ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS X.

1941—1943

SZEGED

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

ACTA
LITTERARUM AC SCIENTIARUM
REGIAE UNIVERSITATIS HUNGARICAE FRANCISCO-JOSEPHINAE

SECTIO:

SCIENTIARUM MATHEMATICARUM.

REDIGUNT:

B. DE KEREKJARTO — F. RIESZ.

TOMUS X.

A M. KIR. FERENCZ JÓZSEF TUDOMÁNYEGYETEM

TUDOMÁNYOS KÖZLEMÉNYEI

MATHEMATIKAI TUDOMÁNYOK.

SZERKESZTIK:

KERÉKJÁRTÓ BÉLA — RIESZ FRIGYES

X. KÖTET.

1941—1943

SZEGED.

A M. Kir. Ferencz József-Tudományegyetem Baráti Egyesületének kiadása.

ACTA
LITTERARUM AC SCIENTIARUM
REGIAE UNIVERSITATIS HUNGARICAE FRANCISCO-JOSEPHINAE

SECTIO :

SCIENTIARUM MATHEMATICARUM.

REDIGUNT :

B. DE KEREKJARTO — F. RIESZ.

TOMUS X.

A M. KIR. FERENCZ JÓZSEF TUDOMÁNYEGYETEM

TUDOMÁNYOS KÖZLEMÉNYEI

MATHEMATIKAI TUDOMÁNYOK.

SZERKESZTIK :

KERÉKJÁRTÓ BÉLA — RIESZ FRIGYES

X. KÖTET.

1941—1943

SZEGED.

A M. Kir. Ferencz József-Tudományegyetem Baráti Egyesületének kiadása.

INDEX — TARTALOM.

Tomus X. — 1941/43. — X. kötet.

BIBLIOGRAPHIE.

Sur la théorie ergodique des espaces abstraits.

Par FRÉDÉRIC RIESZ à Szeged.

Introduction.

Dans une note récente, j'ai démontré quelques théorèmes appartenant à la théorie ergodique[1]). Le présent travail s'attache au dernier de ces théorèmes. Il s'agit dans ce théorème de l'espace fonctionnel L^1, formé des fonctions sommables dans un certain ensemble, soit, pour fixer les idées, dans l'intervalle (0, 1). *Envisageons une transformation linéaire T faisant correspondre à chaque élément f de l'espace L^1 un élément Tf du même espace et supposons que $M_T \leqq 1$, c'est-à-dire que*

$$\int_0^1 |Tf|\,dx \leqq \int_0^1 |f|\,dx$$

pour tous les f. Supposons de plus que pour un certain élément f_1, les moyennes arithmétiques des itérés $f_k = T^{k-1} f_1$,

$$\varphi_n = \frac{1}{n} \sum_1^n f_k,$$

admettent des intégrales dont la continuité absolue est uniforme dans leur ensemble, tel étant le cas, entre autres, lorsque les fonctions $|f_k|$ restent inférieures à la même fonction sommable. Dans ces hypothèses, les moyennes arithmétiques

$$\varphi_{m,n} = \frac{1}{n-m} \sum_{m+1}^n f_k$$

convergent en moyenne d'ordre 1 vers une fonction φ,

$$\int_0^1 |\varphi - \varphi_{m,n}|\,dx \to 0,$$

1) F. RIESZ, Some mean ergodic theorems, *Journal London Math. Society,* 13 (1938), pp. 274—278.

lorsque $n-m\to\infty$ et cette fonction φ est invariante par rapport à la transformation T.

A la fin de la note j'ai promis d'étendre ce théorème, dans un second travail, à un certain type d'espaces abstraits, les espaces (L), envisagés peu avant par M. GARRETT BIRKHOFF[2]), afin de perfectionner le résultat établi par cet auteur. Je me hâte d'avouer que c'est le résultat de M. BIRKHOFF qui m'a inspiré l'idée du théorème que je viens de rappeler. Qu'il nous suffise de dire, pour l'instant, que le type d'espace (L) embrasse entre autres notre espace L^1 et qu'il admet comme celui-ci des transformations linéaires T et alors l'analogue des moyennes φ_n. C'est pour ces moyennes que M. BIRKHOFF établit, sous certaines conditions, une sorte de convergence faible, savoir la convergence des suites numériques $A\varphi_n$ et cela pour toute opération linéaire numérique A.

En étudiant la note de M. BIRKHOFF, il me fallait m'apercevoir et certes l'auteur s'en est aperçu également que, malgré l'admirable finesse du raisonnement, le résultat final attend encore d'être perfectionné en plusieurs directions. Ainsi par exemple il n'y est pas question de l'existence, dans un sens approprié, d'un élément limite φ, tel qu'il intervient dans les théorèmes classiques de MM. G. D. BIRKHOFF et J. DE NEUMANN[3]). Aussi faut il se demander si la convergence faible ne se remplace pas par une sorte de convergence forte? Cherchant à répondre, c'était bien naturel de faire le premier pas en examinant le cas de l'espace fonctionnel L^1 et c'est ainsi que je suis arrivé au théorème ci-dessus.

Ce premier pas fait, je me suis aperçu aussitôt que le passage aux espaces (L), bien que laborieux ici et là, ne présente aucune difficulté sérieuse et qu'on peut le faire par plusieurs voies différentes. Une de ces voies, peut-être la plus courte, consiste à montrer que l'espace (L) ou plutôt un certain sous-espace du même type, le seul qui importe pour le problème actuel, peut être représenté sur l'espace L^1 ou sur un sous-espace de L^1 et cela de sorte que toutes les quantités et toutes les relations qui inter-

---

[2]) G. BIRKHOFF, Dependent probabilities and spaces (L), *Proceedings National Academy USA*, **24** (1938), pp. 154—158.

[3]) Cf. par exemple: G. D. BIRKHOFF and B. O. KOOPMAN, Recent contributions to the ergodic theory, *Proceedings National Academy U. S. A.*, **18** (1932), pp. 279—282.

viennent dans le problème soient conservées. On aboutit à une telle représentation par exemple en combinant quelques détails du présent travail avec un théorème de M. FREUDENTHAL[4]) ou avec un théorème voisin de M. B. DE SZ. NAGY[5]) concernant la réalisation de l'espace H de HILBERT par l'espace fonctionnel L^2 ou par un sous-espace de ce dernier, et cela sous la condition que les fonctions positives correspondent à des éléments de l'espace H fixés d'avance.

On arrive à la même représentation en la faisant asseoir sur la décomposition spectrale des éléments d'un lattis linéaire. Cet ordre d'idées fut découvert indépendamment et esquissé en quelques lignes par M. KAKUTANI dans une note récente[6]). Le passage de (L) à L^1 devient encore plus simple si, au lieu de passer aux fonctions sommables dans un intervalle, on s'arrête dans l'espace à une infinité dénombrable de dimensions.

Dans le présent travail, au lieu de me reporter au théorème particulier concernant l'espace L^1, je préfère de suivre une voie différente, peut-être plus longue mais facile à parcourir puisqu'elle passe, si même à prix d'un détour, par une sorte d'espace de HILBERT, plongé dans l'espace (L) de la même façon que l'espace fonctionnel L^2 l'est dans l'espace L^1. On y parvient en définissant le „carré" et le „produit" des éléments de l'espace (L) comme je l'ai fait il y a quelques années dans la théorie générale des opérations linéaires[7]). En tout cas je pense que la définition nouvelle du carré adoptée dans le présent travail et la facilité avec laquelle on en tire toutes les conséquences, méritent une attention intrinsèque. C'est pour cela que je ne me suis pas efforcé d'épargner quelques détails que l'on aurait pu éluder, par exemple en n'introduisant le carré et le produit que pour les éléments bornés par rapport à une unité préfixée.

---

[4]) H. FREUDENTHAL, Teilweise geordnete Moduln, *Proceedings Academy Amsterdam*, 39 (1936), pp. 641—651.

[5]) B. DE SZ. NAGY, On the set of positive functions in L_2, *Annals of Math.*, 39 (1938), pp. 1—13.

[6]) SH. KAKUTANI, Mean ergodic theorem in abstract (L)-spaces, *Proceedings Imperial Academy Tokyo*, 15 (1929), pp. 121—123.

[7]) F. RIESZ, A lineáris operációk elméletének néhány alapvető fogalomalkotásáról, *Math. és Term.-tud. Értesítő, Budapest*, 56 (1937), pp. 1—46. Une version française, avec quelques modifications, est imprimée dans les *Annals of Math.*, 41 (1940), pp. 173—206; cependant, dans cette version, l'analyse du carré et du produit est remplacée par un calcul fonctionnel général.

Il convient d'ajouter qu'on pourrait aussi faire entrer nos théorèmes comme cas particuliers sous un énoncé de forme générale de M. Yosida[8]) ou sous un autre, encore plus général, établi tout récemment par MM. Alaoglu et Garrett Birkhoff[9]); leur travail arrive au moment où je finis mon manuscrit. Cependant, pour le faire, il faudrait tout d'abord analyser la notion de convergence faible dans l'espace (L) et une telle analyse se base en substance sur les méthodes du présent travail.

Enfin qu'il me soit permis de dire que le délai dans la publication des idées que je vais exposer, est causé par le fait que je les ai réservées pour une conférence que j'étais invité à faire à l'occasion du congrès international qui aurait dû avoir lieu cette année aux Etats-Unis.

I. Les espaces (L).

1. Dans ce premier Chapitre nous énumérons les faits les plus essentiels concernant les lattis linéaires et en particulier les espaces (L), faits dont la plupart se trouvent exposés dans divers travaux sur le sujet[10]).

Nous convenons de désigner les éléments d'un espace abstrait (L) par les lettres f, g, h, φ et ψ.

Les espaces (L) se définissent par les hypothèses suivantes.

1° (L) est une *variété linéaire réelle,* c'est-à-dire qu'il comprend les combinaisons linéaires à coefficients numériques réels de ses éléments, l'addition des éléments et leur multiplication par des facteurs numériques obéissant aux règles ordinaires du Calcul vectoriel.

2° (L) est à la fois un *lattis linéaire,* c'est-à-dire qu'il est partiellement ordonné et cela en considérant comme non-négatifs (ou plus brièvement positifs) certains éléments f, en formule $f \geqq 0$, et en écrivant $f \geqq g$ ou $g \leqq f$ quand $f-g \geqq 0$. On suppose que

a) $f_1 \geqq 0$, $f_2 \geqq 0$ entraînent $f_1+f_2 \geqq 0$;

---

[8]) K. Yosida, Mean ergodic theorem in Banach spaces, *Proceedings Imperial Academy Tokyo,* 14 (1937), pp. 292—294.

[9]) L. Alaoglu and G. Birkhoff, General ergodic theorems, *Annals of Math.,* 41 (1940), pp. 293—309.

[10]) Cf. outre l. c. [2]) et [5]): L. V. Kantorovich, Lineare halbgeordnete Räume, *Recueil math. (Mat. sbornik) Moscou,* 44 (1937), pp. 121—168.

b) $f \geqq 0$, $\lambda \geqq 0$ où λ est un facteur numérique, entraînent $\lambda f \geqq 0$ et inversement $f \geqq 0$, $f \neq 0$ et $\lambda f \geqq 0$ entraînent $\lambda \geqq 0$;

c) pour deux éléments quelconques f_1, f_2 il y a toujours un *plus petit majorant* $g_0 = \sup(f_1, f_2)$, caractérisé par les relations $g_0 \geqq f_1$, $g_0 \geqq f_2$ et $g_0 \leqq g$ pour tout autre g jouissant de la même propriété.

3° (L) est un *espace complet*[11]) dans la métrique définie par une certaine *norme linéaire* $|f|$; cela veut dire que çette norme coïncide, pour f positif, avec une opération linéaire positive fixée If et que, en général, la norme est donnée par la valeur $I(|f|)$ où $|f|$, le *module* de f, est défini par $\sup(f, -f)$.

2. Voici quelques conséquences plus ou moins immédiates de la définition.

Lorsque $f \geqq g$ on a aussi $f+h \geqq g+h$ et $-f \leqq -g$; d'une façon plus générale on a $\lambda f \geqq$ ou $\leqq \lambda g$ suivant que $\lambda \geqq$ ou $\leqq 0$.

Avec une notation évidente, deux éléments quelconques f_1, f_2 admettent un *plus grand minorant* $\inf(f_1, f_2) = -\sup(-f_1, -f_2)$ et l'on a

$$\inf(f_1, f_2) + \sup(f_1, f_2) = f_1 + f_2. \tag{1}$$

En effet, en écrivant φ et ψ pour les deux termes du premier membre, on a $\varphi \leqq f_1$ et $\leqq f_2$, donc aussi $f_1 + f_2 - \varphi \geqq f_2$ et $\geqq f_1$ et alors $f_1 + f_2 - \varphi \geqq \sup(f_1, f_2)$, çe qui donne $f_1 + f_2 \geqq \varphi + \psi$. L'inégalité opposée $f_1 + f_2 \leqq \varphi + \psi$ s'obtient d'une manière analogue en partant de ψ, et les deux inégalités donnent l'égalité (1) qu'il fallait prouver.

3. On a évidemment $|\lambda f| = |\lambda|\,|f|$.

On a par définition $|f| \geqq f$ et $\geqq -f$; en ajoutant et en divisant par 2, on obtient $|f| \geqq 0$. De plus, lorsque $f \geqq 0$, on a aussi $f \geqq -f$ et par conséquent $|f| = f$. L'inverse est évident.

L'inégalité

$$|f_1 + f_2| \leqq |f_1| + |f_2|, \tag{2}$$

qui s'écrit aussi

$$|f| - |g| \leqq |f - g|, \tag{3}$$

---

[11]) Par celą on entend comme d'ordinaire que la condition $|f_m - f_n| \to 0$ $(m, n \to \infty)$ assure l'existence d'un élément limite f, $f_n \to f$, c'est que $|f - f_n| \to 0$. D'ailleurs l'hypothèse *c)* n'est pas essentielle puisqu'elle peut être réalisée toujours en ajoutant à l'espace des éléments idéaux et de plus, nos résultats s'interprètent sans difficulté pour des espaces incomplets.

est valable dans l'espace (L), de même que l'inégalité

(4) $$|f_2-f_1| \geqq \inf(f_2, g) - \inf(f_1, g).$$

Ces inégalités viennent de l'inégalité évidente

$$\sup(f_1+f_2, g_1+g_2) \leqq \sup(f_1, g_1) + \sup(f_2, g_2),$$

la première en y posant $g_1 = -f_1$, $g_2 = -f_2$ et la dernière en remplaçant d'abord f_2, g_1 et g_2 respectivement par f_2-f_1, g et 0, puis en observant que $\sup(f_2-f_1, 0) \leqq |f_2-f_1|$ et enfin en changeant les signes de f_1, f_2 et g pour passer des majorants aux minorants.

4. Des inégalités (2) et (3), en y appliquant l'opération I, on passe immédiatement aux suivantes, concernant la norme:

(5) $$\|f_1+f_2\| \leqq \|f_1\| + \|f_2\|,$$

(6) $$\big|\|f\| - \|g\|\big| \leqq \||f|-|g|\| \leqq \|f-g\|.$$

De cette dernière on conclut que si $f_n \to f$, c'est que $\|f-f_n\| \to 0$, on a aussi $|f_n| \to |f|$ et $\|f_n\| \to \|f\|$. D'ailleurs, à plus forte raison, on a aussi $If_n \to If$.

5. Il nous faudra encore envisager ce qu'on entend par plus petit majorant et plus grand minorant quand, au lieu de deux éléments, il s'agit d'un ensemble $\{f\}$ d'éléments f, fini ou infini. Leur définition est exactement la même que pour deux éléments et leur existence est évidente pour un nombre fini d'éléments.

Avant de passer au cas d'une infinité d'éléments, il sera avantageux de formuler une dernière hypothèse ou plutôt une convention qui ne servira qu'à simplifier le langage. Dans le cas de l'espace L^1 des fonctions sommables, cette convention revient à ne pas regarder comme distinctes deux fonctions qui ne diffèrent que dans un ensemble de mesure nulle. Dans le cas général, *nous convenons d'écrire $f = g$ toujours quand $\|f-g\| = 0$ et $f \geqq 0$ lorsque $|f| = f$ d'après la convention faite* Une telle convention ne modifie pas l'ordre pour les couples d'éléments pour lesquelles l'ordre était déjà fixé; de plus elle ne se fera pas sentir quand il s'agira de transformations linéaires, des éléments „nuls" étant transformés toujours en des éléments „nuls". Notre convention implique entre autres que les suites d'éléments positifs ne peuvent converger que vers des limites positives. En effet, de $0 \leqq f_n \to f$ il s'ensuit par (6) que $f_n = |f_n| \to |f|$, donc $|f| = f$, c'est-à-dire que $f \geqq 0$.

Cela posé, envisageons une *suite croissante* f_n et supposons que la suite numérique If_n soit *bornée.* Alors on voit immédiatement que $|f_n - f_m| = If_n - If_m \to 0$ pour $n > m \to \infty$ et par conséquent f_n *converge vers une limite* f. Comme de plus, pour $n > m$, $f_n - f_m \geqq 0$ et $f_n - f_m \to f - f_m$ avec $n \to \infty$, on a aussi $f - f_m \geqq 0$. Soit d'autre part $g \geqq f_n$ pour tous les n; alors il s'ensuit de la même façon que $g \geqq f$. Ainsi *nous avons montré que*

$$f = \lim f_n = \sup(f_1, f_2, \ldots).$$

6. Soit maintenant $\{f\}$ un ensemble admettant un *majorant* g, c'est-à-dire que $f \leqq g$ pour tous les éléments de l'ensemble. Nous allons montrer que notre ensemble *admet aussi un plus petit majorant.* A cet effet envisageons tous les éléments h du type

$$h = \sup(f_1, \ldots, f_n)$$

où les f_i sont des éléments en nombre quelconque et tirés arbitrairement de l'ensemble. On a évidemment $h \leqq g$, donc $Ih \leqq Ig$ et par conséquent les valeurs Ih admettent une borne supérieure finie μ. Désignons par $h_1, h_2, \ldots$ une suite d'éléments h choisis de sorte que $Ih_i \to \mu$. On pourra aussi supposer que les h_i forment une suite croissante; autrement on n'aurait qu'à remplacer h_i par $\sup(h_1, \ldots, h_i)$. Alors, d'après le N° **5**, les h_i admettent une limite h^* avec $Ih^* = \mu$ et $h^* \leqq g$, quel que soit le majorant g. Soit d'autre part f un élément quelconque de l'ensemble f et posons $h'_i = \sup(h_i, f)$; alors la suite croissante h'_i, du même type que la suite h_i, admettra une limite h' avec $Ih' = \mu$. Or de $h_i \leqq h'_i$ et $f \leqq h'_i$ il s'ensuit immédiatement que $h^* \leqq h'$ et $f \leqq h'$. Donc $|h' - h^*| = Ih' - Ih^* = \mu - \mu = 0$ et par conséquent $h' = h^*$. En somme on a $f \leqq h^* \leqq g$ pour tout élément f et pour tout majorant g de l'ensemble $\{f\}$, c'est-à-dire que h^* fournit précisément $\sup\{f\}$ dont il fallait vérifier l'existence.

II. Carré et produit des éléments.

1. Voici comment on peut introduire le *carré* f^2 d'un élément f et le *produit* $fg = gf$ de deux éléments f et g.

On commence par fixer un élément positif, différent de zéro, que nous désignerons par **1** et qui jouira le rôle d'*unité*. En réalité nous le supprimerons presque toujours dans les formules, en écrivant par exemple $2\lambda f - \lambda^2$ au lieu de $2\lambda f - \lambda^2 \mathbf{1}$.

Nous définissons f^2 par

(7) $$f^2 = \sup(2\lambda f - \lambda^2)$$

où λ varie de $-\infty$ à ∞. Bien entendu on suppose que le second membre ait un sens et s'il en est ainsi on dira que f est *quarrable.* Tel est le cas entre autres lorsque f est *borné* par rapport à l'unité, c'est que $|f| \leqq a = a1$; en effet, dans cette hypothèse,

$$2\lambda f - \lambda^2 \leqq 2|\lambda|\alpha - \lambda^2 \leqq \alpha^2$$

pour toute valeur de λ, ce qui assure l'existence du second membre de (7).

Observons encore que pour f positif, la relation évidente

$$-2\lambda f - \lambda^2 \leqq 2\lambda f - \lambda^2 \qquad (\lambda \geqq 0)$$

permet de se borner, dans la définition qui précède, aux valeurs positives du paramètre λ. Il s'ensuit que *si* $0 \leqq f \leqq g$ *et que* g *est quarrable il en est de même pour* f *et que* $f^2 \leqq g^2$.

Nous allons établir successivement les règles de calcul concernant le carré et le produit des éléments quarrables.

2. *L'élément* f *est quarrable s'il en est ainsi pour son module* $|f|$ *et réciproquement, de plus*

(8) $$f^2 = |f|^2.$$

En effet, on a évidemment

$$f^2 = \sup(2\lambda f - \lambda^2) = \sup[\sup(2\lambda f - \lambda^2, -2\lambda f - \lambda^2)] =$$
$$= \sup[\sup(2|\lambda|f - |\lambda|^2, -2|\lambda|f - |\lambda|^2)] = \sup(2|\lambda||f| - |\lambda|^2) = |f|^2.$$

3. *On a*

(9) $$(\alpha f)^2 = \alpha^2 f^2.$$

Cela vient de la formule

$$2\lambda\alpha f - \lambda^2 = \alpha^2(2\mu f - \mu^2)$$

où l'on a posé $\lambda = \alpha\mu$ et où μ parcourt toutes les valeurs réelles en même temps que λ.

4. *On a*

(10) $$(\alpha f + \beta g)^2 + (\beta f - \alpha g)^2 = (\alpha^2 + \beta^2)(f^2 + g^2)$$

et en particulier

(11) $$(f+g)^2 + (f-g)^2 = 2f^2 + 2g^2.$$

Pour le voir, observons d'abord que le cas général se ramène, en vertu de la formule (9), au cas où $\alpha^2 + \beta^2 = 1$. Cela étant, notre assertion ressort de l'identité

$$[2\lambda(\alpha f+\beta g)-\lambda^2]+[2\mu(\beta f-\alpha g)-\mu^2]=[2\lambda' f-\lambda'^2]+[2\mu' g-\mu'^2]$$

où nous avons posé $\lambda'=\alpha\lambda+\beta\mu$, $\mu'=\beta\lambda-\alpha\mu$.

En effet, λ' et μ' parcourent indépendamment toutes les valeurs réelles en même temps que le font λ et μ et pour obtenir (10) on n'aura qu'à passer des termes entre crochets à leur plus petits majorants.

Notre raisonnement implique que, avec f^2 et g^2, le carré $(\alpha f+\beta g)^2$ existe également ou en d'autres termes, que *les éléments quarrables forment une variété linéaire.*

5. En définissant le produit $fg=gf$ par la formule

$$(f+g)^2=f^2+2fg+g^2, \tag{12}$$

il ressort de (11) que l'on a aussi

$$(f-g)^2=f^2-2fg+g^2 \tag{13}$$

et par conséquent

$$f(-g)=-fg. \tag{14}$$

De plus en remplaçant, dans (10), α par 1 et g par βg et en rappelant la formule (9), on obtient

$$(f+\beta^2 g)^2+\beta^2(f-g)^2=(1+\beta^2)(f^2+\beta^2 g^2)$$

et selon (13) il en ressort que

$$(f+\beta^2 g)^2=f^2+2\beta^2 fg+\beta^4 g^2.$$

En remplaçant g par $-g$ et en appliquant (14), on obtient

$$(f-\beta^2 g)^2=f^2-2\beta^2 fg+\beta^4 g^2$$

et enfin, en écrivant α au lieu de $\pm\beta^2$, les deux dernières formules *se réunissent sous la forme*

$$(f+\alpha g)^2=f^2+2\alpha fg+\alpha^2 g^2 \tag{15}$$

où α désigne un nombre réel quelconque.

6. La formule (15) fournit immédiatement

$$f(\alpha g)=\alpha(fg). \tag{16}$$

De plus en remplaçant, dans la formule (11), f par $f+\frac{1}{2}h$ et g par $g+\frac{1}{2}h$ et en rappelant les formules (12) et (13), on obtient

$$(f+g)h=fh+gh.\text{[12]} \tag{17}$$

[12]) *Cf.* P. Jordan and J. v. Neumann, On inner products in linear metric spaces, *Annals of Math.*, 36 (1936), pp. 719—723.

7. *On a*

(18) $$ff = f^2;\ 0^2 = 0;\ f0 = 0;$$

(19) $$f1 = f.$$

Les relations (18) viennent presque immédiatement des définitions ou encore des formules (9) et (15), en remplaçant g dans cette dernière respectivement par f et par 0.

La relation (19) ressort de l'identité

$$2\lambda(f+1) - \lambda^2 = 2\mu f - \mu^2 + 2f + 1$$

où l'on a posé $\lambda = \mu + 1$. En effet, en passant dans les deux membres au plus petits majorants, on obtient

(20) $$(f+1)^2 = f^2 + 2f + 1.$$

En y posant d'abord $f = 0$, on apprend que

(21) $$1^2 = 1;$$

puis la formule (19) s'ensuit en comparant (20) et (21) avec (12).

8. *Lorsque* $f \geqq 0$, $g \geqq 0$, *on a aussi*

$$fg \geqq 0.$$

Pour le voir, observons d'abord que dans l'hypothèse faite on a $f - g \leqq f + g$, $-f + g \leqq f + g$ et par conséquent $|f - g| \leqq f + g$. De là

$$(f-g)^2 = |f-g|^2 \leqq (f+g)^2$$

et la relation à vérifier s'ensuit immédiatement.

9. *L'hypothèse* $|f_1| \leqq f$, $|g_1| \leqq g$ *entraîne*

(22) $$|f_1 g_1| \leqq fg.$$

En effet

$$f_1 g_1 = fg - \frac{1}{2}(f - f_1)(g + g_1) - \frac{1}{2}(f + f_1)(g - g_1)$$

et par l'hypothèse faite $f \pm f_1 \geqq 0$, $g \pm g_1 \geqq 0$; donc selon ce qui précède,

$$(f \mp f_1)(g \pm g_1) \geqq 0;$$

par conséquent $f_1 g_1 \leqq fg$. Enfin, en remplaçant f_1 par $-f_1$, on obtient $-f_1 g_1 \leqq fg$ et les deux dernières inégalités se réunissent sous la forme (22).

III. Norme quadratique et produit scalaire; l'espace *H*.

1. Outre la norme $|f| = I|f|$ envisagée jusqu'à présent que nous appellerons dans la suite, pour distinguer, la *norme linéaire* de f, introduisons pour les f quarrables, la *norme quadratique*

$$\|f\| = [I(f^2)]^{1/2}.$$

D'après (8) on a évidemment $\| |f| \| = \|f\|$.

Nous envisageons aussi le *produit scalaire*

$$(f, g) = (g, f) = I(fg),$$

se réduisant à $\|f\|^2$ pour $f = g$.

D'après (15), l'expression

$$\|f + \alpha g\|^2 = \|f\|^2 + 2\alpha(f, g) + \alpha^2 \|g\|^2$$

est une forme quadratique positive de la variable α; il s'ensuit que

(23) $$|(f, g)| \leqq \|f\| \, \|g\|.$$

On en déduit aussi, par un raisonnement bien connu, l'inégalité

(24) $$\|f + g\| \leqq \|f\| + \|g\|.$$

2. Lorsque $f \geqq 0$, $g \geqq 0$, on a aussi, d'après **I, 8,**

(25) $$(f, g) \geqq 0.$$

Inversement, la relation (25), *supposée remplie pour tout* $g \geqq 0$, *entraîne* $f \geqq 0$.

Pour le voir, posons $g = |f| - f$; alors notre hypothèse donne

$$I(f|f| - f^2) = (f, g) \geqq 0$$

d'où, d'après (8),

$$I(|f| - f)^2 = I(2f^2 - 2f|f|) \leqq 0;$$

par conséquent $I(|f| - f)^2 = 0$ et $f = |f|$.

3. Soit $f \geqq 0$ et en écrivant $f^{(\nu)} = \inf(f, \nu)$, supposons que $\|f^{(\nu)}\| \leqq A$ pour toute valeur entière positive de ν. Je dis que, *dans cette hypothèse, f est quarrable et* $\|f\| \leqq A$.

En effet, il ressort de l'hypothèse faite que, pour $\nu \to \infty$, les $f^{(\nu)2}$ tendent en croissant vers une limite

$$g = \sup(f^{(\nu)2}),$$

avec $Ig \leqq A^2$.

D'autre part, les $f^{(\nu)}$ eux-mêmes convergent vers f, c'est-à-dire que $|f - f^{(\nu)}| \to 0$. Cela s'ensuit de l'inégalité

$$f = \frac{1}{4\nu} f^2 + \nu - \frac{1}{4\nu}(f - 2\nu)^2 \leqq \frac{1}{4\nu} f^2 + \nu \quad (\nu > 0),$$

dont on conclut que

$$\sup(f, \nu) \leqq \frac{1}{4\nu} f^2 + \nu$$

et alors, d'après (1),

$$0 \leqq f - f^{(\nu)} = \sup(f, \nu) - \nu \leqq \frac{1}{4\nu} f^2,$$

ce qui donne $f^{(\nu)} \to f$.

Enfin

$$2\lambda f^{(\nu)} - \lambda^2 \leqq f^{(\nu)2} \leqq g$$

et pour $\nu \to \infty$ il s'ensuit, d'après ce qui précède, que

$$2\lambda f - \lambda^2 \leqq g;$$

par conséquent f est quarrable et $f^2 \leqq g$, donc aussi $If^2 \leqq Ig \leqq A^2$ ce qu'il fallait démontrer.

4. *Supposons que $f_n \to f$, c'est-à-dire que $|f - f_n| \to 0$ et que $\|f_n\| \leqq A$ pour tous les n; alors f est quarrable et $\|f\| \leqq A$.*

Pour le voir, observons tout d'abord qu'on ne restreint pas la généralité en supposant que $f \geqq 0$, $f_n \geqq 0$. En effet, $\|f\|$ et $\|f_n\|$ ne sont pas changés quand on remplace f et f_n par leurs modules; de plus, grâce à l'inégalité (6), l'hypothèse $f_n \to f$ entraîne $|f_n| \to |f|$.

Supposons donc que $f \geqq 0$ et $f_n \geqq 0$ et posons comme tout à l'heure,

$$f^{(\nu)} = \inf(f, \nu), \; f_n^{(\nu)} = \inf(f_n, \nu).$$

Avec ces notations on a, d'après **II, 9** et la formule (19),

$$|f^{(\nu)2} - f_n^{(\nu)2}| = |(f^{(\nu)} + f_n^{(\nu)})(f^{(\nu)} - f_n^{(\nu)})| \leqq 2\nu |f^{(\nu)} - f_n^{(\nu)}|,$$

donc

$$f_n^{(\nu)2} \to f^{(\nu)2}; \quad A \geqq \|f_n\| \geqq \|f_n^{(\nu)}\| \to \|f^{(\nu)}\|$$

et par conséquent $\|f^{(\nu)}\| \leqq A$. De là notre assertion s'ensuit selon le N° précédent.

5. Envisageons enfin une suite d'éléments quarrables f_n tels que

$$\|f_n - f_m\| \to 0 \qquad (m, n \to \infty).$$

Nous allons montrer qu'*il existe un f quarrable de sorte que*

$$\|f - f_n\| \to 0 \qquad (n \to \infty).$$

Quant à f, celui-ci n'est que la limite des f_n au sens $|f-f_n|\to 0$. En effet, d'après (23),

$$|f_n-f_m|=(|f_n-f_m|,\ 1)\leqq\|f_n-f_m\|\,\|1\|\to 0$$

et cela, avec notre hypothèse 3°, assure la convergence des f_n vers une limite f.

Cela étant, soit ε une quantité positive arbitrairement petite et choisissons m suffisamment grand pour que

$$\|f_n-f_m\|\leqq\varepsilon \tag{26}$$

pour tout $n>m$; alors on aura d'après (24)

$$\|f_n\|\leqq\|f_m\|+\|f_n-f_m\|\leqq\|f_m\|+\varepsilon$$

et par conséquent, d'après le N° précédent, l'élément f est quarrable. De plus comme, pour $n\to\infty$, $f_n-f_m\to f-f_m$, on conclut de (26) par les mêmes raisons

$$\|f-f_m\|\leqq\varepsilon$$

et enfin

$$\|f-f_m\|\to 0 \qquad (m\to\infty),$$

ce qu'il fallait démontrer.

6. En résumé, l'ensemble des éléments quarrables, par rapport à une unité fixée, de l'espace (L) et pour lesquels on définit la distance de f et g par la norme quadratique $\|f-g\|$ ou, ce qui revient au même, le produit scalaire par $(f,g)=I(fg)$, jouit de toutes les propriétés caractéristiques de l'*espace réel de* HILBERT sauf d'une seule; c'est que sa dimension n'est limitée dans aucune direction. Autrement dit, ces éléments forment un *espace euclidien* H à un nombre fini ou à une infinité de dimensions, séparable ou non.

IV. Quelques lemmes.

1. Pour distinguer, convenons d'appeler *convergence en norme quadratique,* brièvement convergence (q) celle qui s'exprime par la relation

$$\|f-f_n\|\to 0$$

comme nous venons de l'envisager et *convergence en norme linéaire* ou convergence (l) celle définie par la relation

$$|f-f_n|\to 0.$$

Un troisième type de convergence dont nous aurons à nous servir, la *convergence faible,* se définit par la relation

$$(f_n, g) \to (f, g),$$

supposée remplie pour tout élément g de l'espace H.

La convergence (q) *implique la convergence* (l) *et aussi la convergence faible;* cela s'ensuit immédiatement de l'inégalité (23).

2. Dans ce qui suit, nous aurons à nous servir de quelques lemmes que nous allons établir. Le premier est bien connu pour l'espace de HILBERT et, à plus forte raison, pour les espaces à dimension finie; le second est compris en substance dans un théorème de MM. BANACH et SAKS[13]); cependant, au lieu de nous reporter à ce dernier, nous préférons de donner une démonstration indépendante.

Lemme 1. *Étant donnée une suite f_n bornée en norme quadratique, $\|f_n\| \leqq A$, on en peut extraire une suite partielle qui converge faiblement.*

Pour le voir, il suffit d'observer que ce fait est bien connu pour l'espace de HILBERT, y inclus les espaces à dimension finie, et que d'autre part, les f_n déterminent un tel espace H_0, qui n'est que le plus petit sous-espace de H qui les comprend. Donc il existe une suite partielle f_{n_k} de sorte que l'on ait

$$(f_{n_k}, g) \to (f, g) \tag{27}$$

pour un certain f appartenant à H_0 et pour tous les g dans H_0. Or il est évident que (27) est vrai aussi pour tous les g dans H qui sont orthogonaux à H_0. Comme enfin on sait que chaque élément de H est la somme de deux composants, l'un dans H_0 et l'autre orthogonal à H_0, la relation (27) est valable pour tout g dans H.

3. Lemme 2. *Lorsque la suite f_n converge faiblement vers l'élément f, ce dernier peut être approché indéfiniment, au sens de la convergence* (q), *par des moyennes arithmétiques du type général des éléments f_n, c'est-à-dire par des combinaisons linéaires*

$$g = \sum_r^s c_k f_k,$$

formées avec des $c_k \geqq 0$, de somme 1. De plus les indices r peuvent être choisis aussi élevés qu'on voudra[14]).

[13]) S. BANACH et S. SAKS, Sur la convergence forte dans les espaces L^p, *Studia Math.*, 2 (1930), pp. 51—57.

[14]) Dans la démonstration du théorème 1, ce lemme peut être remplacé par l'énoncé correspondant moins précis, d'ailleurs bien connu pour l'espace de HILBERT, dans lequel rien n'est exigé des coefficients c_k.

En effet, on peut tout d'abord, sans restreindre la généralité, se borner au cas où $r=1$; autrement on n'aurait qu'à envisager la suite $f_r, f_{r+1}, \ldots$

Cela étant, soit G l'ensemble des éléments g du type envisagé et soit $\bar{G}$ sa fermeture au sens de la norme quadratique. Ce qu'il faut montrer, c'est que f appartient à cette fermeture $\bar{G}$. Or on sait qu'il existe dans $\bar{G}$ un élément $g=g^*$ qui rend minimum la distance $\|f-g\|$ de f et des éléments de $\bar{G}$.[15]) Je dis que $f=g^*$. Pour le montrer, observons que les éléments f_n, donc aussi les éléments

$$g^*+\lambda(f_n-g^*)=(1-\lambda)g^*+\lambda f_n \qquad (0\leqq\lambda\leqq 1)$$

appartiennent à l'ensemble convexe $\bar{G}$. Par conséquent on a, pour $0\leqq\lambda\leqq 1$,

$$\|f-g^*-\lambda(f_n-g^*)\|\geqq\|f-g^*\|$$

et en développant les carrés des deux membres on obtient

$$-2\lambda(f-g^*, f_n-g^*)+\lambda^2\|f_n-g^*\|^2\geqq 0.$$

De là, en divisant par 2λ et en posant $\lambda=0$, il s'ensuit que

$$(f-g^*, f_n-g^*)\leqq 0$$

et enfin, en vertu de la convergence faible des f_n vers f,

$$\|f-g^*\|^2=(f-g^*, f-g^*)\leqq 0,$$

donc $\|f-g^*\|=0$, ce qui achève la démonstration.

4. Lemme 3. *Lorsque la suite f_n converge faiblement vers f et que $f_n\geqq 0$, on a aussi $f\geqq 0$.*

En effet, dans l'hypothèse faite on a, d'après (25), pour tous les $g\geqq 0$,

$$(f_n, g)\geqq 0$$

et par conséquent

$$(f, g)=\lim(f_n, g)\geqq 0$$

ce qui, d'après **III, 2,** entraîne que $f\geqq 0$.

V. Les théorèmes.

1. Après ces préparatifs nous allons démontrer les deux théorèmes ergodiques que nous avons en vue. Quoique le premier ne soit qu'un cas particulier du second, nous préférons de nous en occuper d'abord séparément.

---

[15]) *Cf.* l. c. [5]), en particulier p. 5 en bas.

Théorème 1. *Soit donnée, dans l'espace abstrait* (L), *une transformation linéaire* T *pour laquelle* $M_T \leqq 1$, *c'est que* $|Tf| \leqq |f|$ *pour tous les* f,[16]) *et supposons que pour un certain élément* f_1, *cet élément et tous ses itérés* $f_n = T^{n-1} f_1$ *ou plus généralement, que les moyennes arithmétiques*

$$\varphi_n = \frac{1}{n} \sum_1^n f_k$$

restent inférieures en module à un élément fixe[17]).

Dans ces conditions, les moyennes arithmétiques

$$\varphi_{m,n} = \frac{1}{n-m} \sum_{m+1}^{n} f_k$$

convergent en norme linéaire, pour $n - m \to \infty$, *vers un élément déterminé* φ, *invariant par rapport à la transformation* T.

Pour le voir, on construit l'espace H en convenant de choisir pour 1 l'élément fixe qui figure dans notre théorème; alors on aura évidemment $\|\varphi_n\| \leqq |1|^{1/2}$. Par conséquent, grâce au lemme 1 du Chapitre précédent, la suite φ_n donne lieu à une suite partielle qui converge faiblement vers un élément φ. D'après le lemme 2, φ peut être approché indéfiniment, au sens de la convergence en norme quadratique, par des éléments g du type

$$(28) \qquad g = \sum_r^s c_k \varphi_k \qquad (c_k \geqq 0, \ \sum c_k = 1),$$

avec r aussi grand qu'on voudra. A plus forte raison, la suite des g et alors celle des transformés Tg converge vers φ aussi en norme linéaire. Or l'équation

$$\varphi_n - T\varphi_n = \frac{1}{n}(f_1 - f_{n+1})$$

donne

$$|g - Tg| = \left| \sum_r^s \frac{c_k}{k}(f_1 - f_{k+1}) \right| \leqq \frac{2}{r} |f_1| \to 0$$

pour $r \to \infty$. De là on conclut que les limites respectives φ et $T\varphi$ de g et Tg coïncident, c'est-à-dire que φ est invariant par rapport à la transformation T. De plus, en introduisant au lieu des φ_n leurs

[16]) Observons que l'hypothèse $M_T \leqq 1$ n'intervient que pour assurer que la suite M_{T^n} reste bornée et que, par conséquent, on pourrait la remplacer, ici de même que dans l'énoncé du théorème 2, par cette dernière hypothèse.

[17]) Autrement dit la suite $|\varphi_n|$ admet un majorant ou encore la suite φ_n admet un majorant et un minorant.

expressions par les f_k, on obtient

$$g = \sum_1^s c'_k f_k \qquad (\sum c'_k = 1)$$

et par conséquent on pourra poser

$$\varphi - f_1 = \sum_2^s c'_k (f_k - f_1) + h = \sum_2^s c''_k (f_k - f_{k-1}) + h,$$

où $|h| \leqq \varepsilon$, avec ε arbitrairement petit. De là, en appliquant aux deux membres successivement les transformations $T^m, \ldots, T^{n-1}$ et en formant les moyennes arithmétiques, on obtient

$$\varphi - \varphi_{m,n} = \frac{1}{n-m} \sum_{k=2}^s c''_k (f_{k+n-1} - f_{k+m-1}) + \frac{1}{n-m} \sum_m^{n-1} T^k h,$$

et il s'ensuit que

$$|\varphi - \varphi_{m,n}| \leqq \frac{2C}{n-m} |f_1| + \varepsilon \qquad \left(C = \sum_2^s |c''_k|\right)$$

et que de cette sorte

$$|\varphi - \varphi_{m,n}| \leqq 2\varepsilon$$

pour $n-m$ suffisamment grand, ce qu'il fallait prouver.

2. Avant de formuler le théorème 2, convenons de dire que la suite f_n est *presque bornée* lorsqu'il y a un élément $g \geqq 0$ tel que, pour tout $\varepsilon > 0$, on peut choisir les constantes $\alpha > 0$, $\beta > 0$ de sorte que l'on ait, pour tout n,

$$I \sup (f_n, -\alpha g) \leqq I f_n + \varepsilon, \quad I \inf (f_n, \beta g) \geqq I f_n - \varepsilon.$$

Évidemment, sans restreindre la généralité on pourrait aussi supposer que $\alpha = \beta$.[18])

Dans le cas de l'espace L^1, le fait supposé est équivalent à ce que les intégrales des fonctions $f_n(x)$ sont absolument continues uniformément dans leur ensemble.

Observons que, en même temps que la suite f_n, la suite des moyennes arithmétiques φ_n est aussi presque bornée, ce qui ressort des inégalités évidentes

$$\sup (\varphi_n, -\alpha g) \leqq \frac{1}{n} \sum_1^n \sup (f_k, -\alpha g);$$

$$\inf (\varphi_n, \beta g) \geqq \frac{1}{n} \sum_1^n \inf (f_k, \beta g).$$

---

[18]) On pourrait aussi faire entrer dans la définition les $|f_n|$ au lieu des f_n et remplacer les deux inégalités par une seule.

Cela étant, notre théorème 1 se généralise comme il suit.

Théorème 2. *Soit donnée, dans l'espace abstrait (L), une transformation linéaire T telle que $M_T \leqq 1$ et supposons qu'un certain élément f_1 et ses itérés $f_n = T^{n-1} f_1$ ou plus généralement, que les moyennes arithmétiques φ_n des f_n forment une suite presque bornée.*

Dans ces conditions, les moyennes arithmétiques $\varphi_{m,n}$ convergent en norme linéaire, pour $n-m \to \infty$, vers un élément déterminé φ, invariant par rapport à la transformation T.

La démonstration du théorème ne diffère de celle du théorème 1 qu'en un seul point. C'est que la construction de la limite φ et son approximation par des éléments g du type (28) exige des considérations plus laborieuses. Mais à partir de ce point, la démonstration s'achève exactement comme celle du théorème 1 et on n'a pas besoin de la répéter.

Pour établir le point en question, choisissons pour unité 1 l'élément g par rapport auquel les φ_n sont presque bornés et envisageons la suite des $\varphi_{n;\,\alpha,\,\beta}$, définis par

$$\varphi_{n;\,\alpha,\,\beta} = \sup(\varphi_n, -\alpha) + \inf(\varphi_n, \beta) - \varphi_n, \tag{29}$$

avec $\alpha, \beta \geqq 0$ donnés. En se servant de l'identité (1), on obtient immédiatement que

$$\varphi_{n;\,\alpha,\,\beta} = -\alpha + \inf(\varphi_n, \beta) - \inf(\varphi_n, -\alpha) \geqq -\alpha$$

et de même

$$\varphi_{n;\,\alpha,\,\beta} \leqq \beta.$$

Il s'ensuit évidemment que la suite est aussi bornée en norme quadratique et que par conséquent, grâce au lemme 1, elle admet une suite partielle qui converge faiblement vers un élément $\varphi^{(\alpha,\,\beta)}$. Enfin, en faisant varier α et β et en se servant du procédé diagonal bien connu, on arrive à une suite n_k telle que les suites $\varphi_{n_k;\,\alpha,\,\beta}$ convergent faiblement, pour $k \to \infty$, vers des limites respectives $\varphi^{(\alpha,\,\beta)}$ quels que soient les entiers α, β.

Pour passer des $\varphi^{(\alpha,\,\beta)}$ à l'élément φ qui figure dans notre théorème, observons tout d'abord que la formule (29) par laquelle nous venons de définir les éléments $\varphi_{n;\,\alpha,\,\beta}$, montre nettement que $\varphi_{n;\,\alpha,\,\beta}$ croît avec α décroissant et aussi avec β croissant. Grâce au lemme 3, ce fait subsiste pour $\varphi^{(\alpha,\,\beta)}$. Pour aller plus loin, il nous faut envisager la valeur $I\varphi^{(\alpha,\,\beta)}$. Nous allons montrer qu'elle est comprise entre deux bornes ne dépendant ni de α ni de β.

Nous commençons par calculer de telles bornes pour les valeurs $I\varphi_{n;\alpha,\beta}$, bornes qui seront indépendantes aussi de l'indice n.

Comme, par hypothèse, la suite φ_n est presque bornée par rapport à $g=1$, on peut choisir $\alpha_0>0$ et $\beta_0>0$ de sorte que les différences $I(\sup(\varphi_n, -\alpha_0))-I\varphi_n$ et $I\varphi_n - I(\inf(\varphi_n, \beta_0))$ soient inférieures à l'unité. D'autre part, la formule (29) fournit immédiatement

$$\varphi_{n;\alpha,\beta} = \varphi_{n;\alpha,\beta_0} + \inf(\varphi_n, \beta) - \inf(\varphi_n, \beta_0) \leqq \beta_0 + \varphi_n - \inf(\varphi_n, \beta_0)$$

et de là il ressort que

$$I\varphi_{n;\alpha,\beta} \leqq \beta_0 |\mathbf{1}| + 1.$$

Un calcul analogue donne

$$I\varphi_{n;\alpha,\beta} \geqq -\alpha_0 |\mathbf{1}| - 1.$$

C'est-à-dire que $I\varphi_{n;\alpha,\beta}$ est compris entre les deux bornes indiquées, ne dépendant ni de α, ni de β, ni de n. Enfin, comme

$$I\varphi_{n;\alpha,\beta} = (\varphi_{n;\alpha,\beta}, \mathbf{1}) \to (\varphi^{(\alpha,\beta)}, \mathbf{1}) = I\varphi^{(\alpha,\beta)}$$

à cause de la convergence faible, la quantité $I\varphi^{(\alpha,\beta)}$ sera comprise entre les mêmes bornes et cela pour tous les α et β.

Cela établi, pour passer à φ, on n'a qu'à fixer d'abord α, faire β aller à l'infini et envisager la suite croissante $\varphi^{(\alpha,\beta)}$. D'après **I, 5,** cette suite converge en norme linéaire vers un élément $\varphi^{(\alpha)} = \sup_\beta \{\varphi^{(\alpha,\beta)}\}$. Les valeurs $I\varphi^{(\alpha)}$, limites de $I\varphi^{(\alpha,\beta)}$, restent aussi entre les bornes indiquées et par conséquent $\varphi^{(\alpha)}$ qui va en décroissant, converge en norme linéaire, pour α infini, vers une limite φ.

Il nous reste d'approcher cette limite φ par des éléments g du type (28). Dans ce but, après avoir donné $\varepsilon>0$, choisissons α et β de sorte que l'on ait à la fois

$$(30) \qquad I\varphi^{(\alpha)} - I\varphi \leqq \frac{\varepsilon}{5}; \quad I\varphi^{(\alpha)} - I\varphi^{(\alpha,\beta)} \leqq \frac{\varepsilon}{5}$$

et

$$I(\sup(\varphi_n, -\alpha)) - I\varphi_n \leqq \frac{\varepsilon}{5}; \quad I\varphi_n - I(\inf(\varphi_n, \beta)) \leqq \frac{\varepsilon}{5},$$

donc aussi

$$(31) \quad |\varphi_n - \varphi_{n;\alpha,\beta}| = |\varphi_n - \sup(\varphi_n, -\alpha) + \varphi_n - \inf(\varphi_n, \beta)| \leqq \frac{2\varepsilon}{5}$$

et cela pour tous les n. Enfin, la suite $\varphi_{n_k;\alpha,\beta}$ convergeant faiblement vers $\varphi^{(\alpha,\beta)}$, il existe, d'après le lemme 2, des combinaisons linéaires

$$g = \sum_{r}^{s'} c_n \varphi_{n;\,\alpha,\beta} \tag{32}$$

avec $c_n \geqq 0$, $\Sigma c_n = 1$ et r arbitrairement élevé, de sorte que

$$|\varphi^{(\alpha,\beta)} - g| \leqq \frac{\varepsilon}{5}. \tag{33}$$

En combinant les formules (30)—(33) on obtient que

$$\left| \varphi - \sum_{r}^{s} c_n \varphi_n \right| \leqq \varepsilon$$

et c'est précisément le point qu'il nous restait d'établir. A partir de là, comme nous l'avons déjà dit, la démonstration s'achève mot à mot comme celle du théorème 1.

(Reçu le 27 mai 1940)

Correction. Page 16, ligne 13 de bas, *au lieu de* converge vers φ *lire* convergent respectivement vers φ et $T\varphi$.

Sur les groupes transitifs de la droite.

Par B. DE KERÉKJÁRTÓ à Budapest.

A M. L. E. J. Brouwer
en témoignage de haute estime.

Introduction.

Un groupe G de transformations topologiques d'une variété V en elle-même est appelé *n-uplement transitif* sur V, si à deux systèmes quelconques de n points de $V: (A_1, A_2, \ldots, A_n)$ et $(A'_1, A'_2, \ldots, A'_n)$ correspond une transformation T du groupe G et une seule qui change le premier système en le second (de telle manière qu'à A_ν correspond le point image A'_ν). Le groupe n-uplement transitif G est dit *continu* si la transformation T varie continuement avec le système $(A'_1, A'_2, \ldots, A'_n)$.

Dans le § 1 du présent mémoire je démontre que *tout groupe simplement transitif de la droite, de même que tout groupe n-uplement transitif de la droite, en tant qu'un tel groupe existe, est continu.* Cependant la droite n'admet que des groupes simplement ou doublement transitifs, et la droite projective des groupes simplement ou triplement transitifs.

M. BROUWER a démontré dans ses travaux importants sur les groupes continus[1]) que *les groupes continus simplement, doublement et triplement transitifs de la droite sont homéomorphes respectivement aux groupes suivants: groupe des translations, groupe des similitudes et groupe des homographies de la droite.*

Dans le § 2 je reproduis la démonstration de M. BROUWER de la caractérisation du groupe simplement transitif. Concernant

[1]) L. E. J. BROUWER, Die Theorie der endlichen kontinuierlichen Gruppen unabhängig von den Axiomen von LIE, I., *Math. Annalen,* **67** (1909), p. 246—267.

les groupes doublement ou triplement transitifs, M. BROUWER a démontré que les transformations de ces groupes s'expriment par des fonctions dérivables du paramètre canonique d'un sous-groupe d'ordre 1 y contenu. Par la solution des équations différentielles respectives, il a obtenu la caractérisation de ces groupes.

Dans les §§ 3, 4 de ce mémoire je donne des démonstrations nouvelles et élémentaires des théorèmes de M. BROUWER en utilisant des méthodes qui, dans un mémoire antérieur[2]), m'ont permis de caractériser topologiquement le groupe homographique de la sphère.

§ 1. La continuité des groupes transitifs de la droite.

1. 1. *Tout groupe simplement transitif de transformations topologiques de la droite en elle-même est continu.*

Aucune transformation du groupe n'admet un point invariant, sauf l'identité; par suite, toute transformation du groupe conserve le sens de la droite.

Conformément à un sens positif de la droite, nous associons à tout couple de points distincts A, B la relation $A < B$ ou $B < A$, suivant que le segment AB est de sens positif ou négatif.

Soit T une transformation du groupe. Si pour un point A, la relation $A < T(A)$ est valable, on a, pour tout point B, la relation $B < T(B)$; car autrement le segment AB serait transformé par T en une partie de lui-même, et la transformation T aurait donc un point invariant sur le segment AB.

Si pour deux points quelconques A et B la relation $A < B$ est valable, on a, pour toute transformation T du groupe, la relation $T(A) < T(B)$; cette proposition est une conséquence immédiate du fait que T conserve le sens de la droite.

Soit O un point quelconque, et soit $A_1, A_2, \ldots$ une suite monotone de points convergeant vers un point A; supposons, pour fixer les idées, que $A_1 < A_2 < \ldots < A$. Désignons par T et par T_ν les transformations du groupe qui changent le point O en A et en A_ν respectivement. Nous allons démontrer que la suite des transformations $T_1, T_2, \ldots$ converge vers T.

---

[2]) B. DE KERÉKJÁRTÓ, Sur le caractère topologique du groupe homographique de la sphère. A paraître dans le *Journal de Mathématiques pures et appliquées.*

Soit Q un point quelconque différent de O; désignons par $B = T(Q)$ et $B_\nu = T_\nu(Q)$ ses images obtenues par T et par T_ν. De la relation $A_\nu < A_{\nu+1}$, on conclut que $B_\nu < B_{\nu+1}$; en effet la transformation $T_\nu^{-1} T_{\nu+1}$ du groupe transforme A_ν en $A_{\nu+1}$ et B_ν en $B_{\nu+1}$. D'une façon similaire on obtient la relation $B_\nu < B$ valable pour tout ν.

Si la suite des points $B_1, B_2, \ldots$ ne converge pas vers B, désignons par B' la limite de cette suite monotone et bornée; on a la relation $B' < B$. Soit T' la transformation du groupe qui change le point Q en B'; désignons par $A' = T'(O)$ l'image du point O obtenue par T'. Des relations $B' < B$ et $B_\nu < B'$ on déduit les relations $A' < A$ et $A_\nu < A'$. On aurait donc, pour tout ν, la relation $A_\nu < A' < A$; c'est en contradiction avec notre hypothèse: $A = \lim A_\nu$.

Nous avons ainsi démontré que, pour tout point Q, la suite des points $T_\nu(Q)$ converge vers le point $T(Q)$. Il résulte des théorèmes connus de la théorie des fonctions monotones que les transformations T_ν convergent uniformément vers la transformation T, c'est-à-dire que, dans une métrique bornée de la droite, à toute quantité positive ε correspond un indice ν_0 tel que, pour tout $\nu > \nu_0$, les transformations T_ν et T diffèrent de moins de ε.

1. 2. *Tout groupe doublement transitif de transformations topologiques de la droite en elle-même est continu.*

Il faut montrer qu'à un couple quelconque de points A, B et à tout nombre positif ε correspond un nombre positif δ tel que toute transformation du groupe donné G qui déplace chacun des points A, B de distances inférieures à δ diffère de l'identité de moins de ε.

Les transformations du groupe G qui laissent invariant le point A et conservent le sens de la droite forment un sous-groupe g_A de G lequel est *simplement transitif* sur la demi-droite AB. En vertu de la proposition 1. 1, on peut déterminer un nombre positif η tel que toute transformation de g_A qui déplace B de moins de η diffère de l'identité de moins de $\varepsilon/2$; il est évident que $\eta < \varepsilon$. D'une façon similaire, on peut déterminer un nombre positif δ $(< \eta/2)$ tel que toute transformation de G qui laisse B invariant et déplace A de distance inférieure à δ diffère de l'identité de moins de $\eta/2$.

Ceci posé, soient A' et B' deux points arbitraires tels que les distances (A, A') et (B, B') soient inférieures à δ. Désignons par T_1 la transformation de G qui laisse B invariant et change A en A'; comme $(A, A') < \delta$, il résulte de nos conditions que la transformation T_1 diffère de l'identité de moins de $\eta/2$. Soit $B'' = T_1^{-1}(B')$; la distance (B', B'') est plus petite que $\eta/2$, et $(B, B') < \delta < \eta/2$, d'où $(B, B'') < \eta$. Il en résulte que la transformation T_2 de G qui laisse le point A invariant et change B en B'' diffère de l'identité de moins de $\varepsilon/2$. Le produit $T_2 T_1$ des transformations T_2 et T_1 change le point A en A' et B en B'; elle diffère de l'identité de moins de $(\eta + \varepsilon)/2 < \varepsilon$.

1. 3. *Il n'y a aucun groupe n-uplement transitif de transformations topologiques de la droite en elle-même, pour $n > 2$.*

Les transformations d'un groupe n-uplement transitif, pour $n > 2$, qui admettent deux points invariants A et B formeraient un sous-groupe qui serait encore transitif sur la droite privée des points A et B. Mais cela est impossible parce que toute transformation topologique de la droite en elle-même qui laisse invariants les points A et B change les points intérieurs au segment AB entre eux.

Concernant la *droite projective* (qui est homéomorphe à un cercle) on déduit par des raisonnements analogues les propositions suivantes:

1. 4. *Tout groupe simplement ou triplement transitif de transformations topologiques de la droite projective en elle-même est continu.*

1. 5. *Il n'y a aucun groupe n-uplement transitif de transformations topologiques de la droite projective en elle-même, pour $n > 3$.*

Nous allons démontrer la proposition suivante:

1. 6. *Il n'y a aucun groupe doublement transitif de transformations topologiques de la droite projective en elle-même.*

Si G est un groupe doublement transitif de transformations topologiques de la droite projective en elle-même, l'identité est la seule transformation de G qui admet deux points invariants. Toute transformation de G conserve, par conséquent, le sens de la droite, parce qu'une transformation topologique de la droite projective changeant le sens en le sens opposé admet deux points invariants.

Soient A et B deux points quelconques; il y a dans le groupe G une transformation et une seule qui change le couple de points A, B en B, A; désignons cette transformation par σ^{AB} Le carré de σ^{AB} est l'identité, parce qu'il admet deux points invariants, A et B; la transformation σ^{AB} est donc *involutive.*

Soit C un point différent de A et de B et désignons par σ^{AC} la transformation involutive du groupe G qui échange les points A et C entre eux. Le produit $\tau = \sigma^{AB}\sigma^{AC}$ des transformations σ^{AB} et σ^{AC} change B en C et A en le point $B' = \sigma^{AC}(B)$. Les deux segments de la droite projective déterminés par les points A, C, lesquels nous désignerons par AC et par CA, sont échangés entre eux par la transformation σ^{AC}; nous concluons de là que le couple de points A, C sépare les points B et B'. L'un des segments de la droite projective déterminés par les points A, B contient donc les points B' et C; désignons ce segment par AB. La transformation τ change le segment AB en le segment $B'C$ qui est une partie du segment AB; il résulte que τ admet un point invariant sur le segment AB. Similairement l'inverse de τ change le segment CB' en le segment BA contenu dans l'intérieur de CB'; la transformation τ admet, par suite, un point invariant sur le segment BA. La transformation τ admettant deux points invariants devrait être l'identité; mais, d'autre part, τ change le point B en le point C différent de B. Nous sommes aboutis à une contradiction.

§ 2. Les groupes simplement transitifs de la droite.

Théorème. *Tout groupe simplement transitif de transformations topologiques de la droite en elle-même est homéomorphe au groupe des translations.*

La démonstration suivante est due à M. Brouwer; nous la reproduisons ici en vue de comparaison avec la méthode du § 3.

Soit O un point quelconque de la droite et soit T une transformation du groupe donné, différente de l'identité. Désignons par $A_1 = T(O)$, $A_2 = T^2(O), \ldots, A_n = T^n(O), \ldots$ les images successives de O obtenues par l'itération de T, et par $A_{-1} = T^{-1}(O)$, $A_{-2} = T^{-2}(O), \ldots$ les images de O obtenues par l'itération de l'inverse de T.

La transformation T n'admet aucun point invariant, elle conserve donc le sens de la droite; par suite, les segments OA_1 et A_1A_2 ont même sens. Il en résulte que l'ordre des points A_n sur la droite est le même que l'ordre naturel des entiers n qui figurent dans leurs indices.

La suite des points $A_1, A_2, \ldots$ *est divergente.* Autrement le point-limite B de cette suite monotone serait un point invariant de la transformation T; ce serait une contradiction avec nos hypothèses.

Soit X un point variable du segment OA_1 et soit S la transformation du groupe qui change O en X; désignons par Y l'image de X obtenue par la même transformation S. Si X varie d'une façon continue et monotone à partir de O jusqu'à A_1, le point correspondant Y varie d'une façon continue et monotone à partir de O jusqu'à A_2. Il y a donc un point X et un seul du segment OA_1 pour lequel Y coïncide avec le point A_1. Nous désignons ce point par $A_{1/2}$, et la transformation qui change O en $A_{1/2}$ par $T^{1/2}$. D'une façon analogue, nous définissons successivement les points $A_{1/4}, A_{1/8}, \ldots$ et les transformations $T^{1/4}, T^{1/8}, \ldots$ leur correspondant.

La suite des points $A_{1/2}, A_{1/4}, \ldots$ *converge vers* O. En effet, si A est un point quelconque entre O et $A_{1/2^n}$, et si S est la transformation du groupe qui change O en A, le point $S(A)$ se trouve entre O et $T^{1/2^n}(A)$, et ce dernier point se trouve entre O et $T^{1/2^n}(A_{1/2^n}) = A_{1/2^{n-1}}$. Nous en concluons que, si A est la limite de la suite $A_{1/2^n}$, $S(A)$ est la limite de la suite $A_{1/2^{n-1}}$, d'où $S(A) = A$. De cette relation découle que S est l'identité, et $A = S(O) = O$.

Attribuons à l'élément $T^{p/2^n} = (T^{1/2^n})^p$ le paramètre $\xi = p/2^n$ et au point $T^{p/2^n}(O)$ la coordonnée $x = p/2^n$. L'ordre des coordonnées x est le même que l'ordre des points de la droite leur correspondant. Le produit des transformations de paramètres ξ et ξ' est la transformation qui correspond à la valeur $\xi + \xi'$ du paramètre. L'attribution du paramètre ξ et de la coordonnée x s'étend par continuité à tous les points de la droite. Avec ce système de coordonnée et de paramètre, toute transformation du groupe s'exprime par la formule: $x' = x + \xi$.

§ 3. Les groupes doublement transitifs de la droite.

Théorème. *Tout groupe doublement transitif de transformations topologiques de la droite en elle-même est homéomorphe au groupe des similitudes.*

Soit G un groupe doublement transitif de transformations topologiques de la droite en elle-même.

3. 1. *A un couple quelconque de points A, B correspond une transformation involutive du groupe donné G et une seule qui échange entre-eux les points A et B. Cette transformation admet un seul point invariant C, appartenant à l'intervalle AB.*

Il y a une transformation σ dans le groupe G et une seule qui transforme le couple A, B en le couple B, A. Le carré de σ admet deux points invariants, A et B, il est donc l'identité : $\sigma^2 = I$; σ est, par suite, involutive.

Si le point variable X décrit l'intervalle AB, l'image $X' = \sigma(X)$ de X obtenue par σ décrit le même intervalle dans le sens opposé. Il y a un point C et un seul, appartenant à l'intervalle AB, pour lequel $C = \sigma(C)$. Nous appelons *C milieu du segment AB.* Nous venons de montrer que *tout segment a un milieu et un seul.*

3. 2. *A tout point P de la droite correspond une transformation involutive de G et une seule admettant le point invariant P.*

Soit A, B un couple quelconque, soit $\bar{\sigma}$ la transformation involutive qui échange entre eux les points A et B, et soit C le point invariant de $\bar{\sigma}$ (voir 3. 1). Désignons par T une transformation du groupe G qui change le point C en P; la transformée de $\bar{\sigma}$ par T, c'est-à-dire la transformation $\sigma = T^{-1}\bar{\sigma}T$ est involutive et admet le point invariant $T(C) = P$. Il y a, par conséquent, au moins une transformation involutive dans G qui admet le point invariant P.

Si σ' était une autre transformation involutive de G qui admet le point invariant P, désignons par $B' = \sigma'(A)$ l'image d'un point arbitraire A obtenue par σ', et par $A' = \sigma(B')$ l'image de B' obtenue par σ. Soit T la transformation de G qui change A en A' et laisse B' invariant. La transformée de σ' par $T: \sigma'' = T^{-1}\sigma' T$ échange entre eux les points A' et B' de même que σ; il résulte de là que σ est identique à σ''. Par suite le point invariant $P' = T(P)$ de σ'' est identique à P. La transformation T admet deux points invariants B' et P, donc $T = I$, et $\sigma = \sigma'' = T^{-1}\sigma' T = \sigma'$.

Nous désignerons par σ_P la transformation involutive du groupe G dont P est le point invariant; nous l'appellerons *symétrie* par rapport au point P.

3. 3. *Si une transformation T de G change les points A, B en A', B', elle change le milieu C du segment AB en le milieu C' du segment A'B'.*

Soit en effet σ_C la symétrie par rapport à C. Sa transformée

$T^{-1}\sigma_C T$ est la symétrie par rapport à $C' = T(C)$; elle échange entre eux les points A' et B'.

Désignons par O un point quelconque de la droite. Les transformations du groupe G qui laissent invariant le point O forment un sous-groupe que nous désignons par G_O. Les transformations de G_O qui conservent le sens de la droite forment un sous-groupe g_O de G_O qui est *simplement transitif* sur l'une quelconque des demi-droites déterminées par le point O. D'après le théorème démontré dans le § **2**, le groupe g_O est homéomorphe au groupe des translations $y' = y + y_0$; en introduisant un autre paramètre par les formules: $X = e^y$, $X_0 = e^{y_0}$, on peut exprimer g_O par la formule: $X' = XX_0$ $(X_0 > 0)$. On conclut de là, en particulier, que *le groupe g_O est commutatif.*

Par multiplication des éléments du groupe g_O par la symétrie σ_O, nous obtenons toutes les transformations de G_O n'appartenant pas à g_O. *Le groupe g_O est échangeable avec σ_O*, c'est-à-dire que $\sigma_O \mu = \mu \sigma_O$, pour tout élément μ de g_O. Soit en effet A un point quelconque, $A' = \sigma_O(A)$, $B = \mu(A)$, $B' = \mu(A')$. D'après 3. 3, O est le milieu du segment BB', d'où $\sigma_O(B) = B'$. Pour un point arbitraire A, on a donc les relations $\sigma_O \mu(A) = \mu \sigma_O(A)$. Nous avons ainsi démontré la proposition suivante:

3. 4. *Les transformations du groupe G qui laissent invariant le point O forment un sous-groupe commutatif G_O. Le groupe G_O est formé par le sous-groupe g_O de G_O, conservant le sens de la droite, et par le produit de g_O par la symétrie σ_O relative au point O.*

3. 5. *Le produit de deux symétries n'admet aucun point invariant.*

Soient en effet O et A deux points distincts quelconques; si le produit $\sigma_O \sigma_A$ des symétries σ_O et σ_A laisse le point P invariant, soit $P' = \sigma_O(P)$, on a donc $\sigma_A(P') = P$. Les points P et P' sont distincts. Comme σ_O et σ_A sont involutives, on a aussi $\sigma_O(P') = P$ et $\sigma_A(P) = P'$; par suite, le point P' est un autre point invariant de $\sigma_O \sigma_A$. Il en résulte que $\sigma_O \sigma_A$ est l'identité ce qui est une contradiction.

3. 6. *Soit $\tau = \sigma_O \sigma_A$ le produit de deux symétries. Les images successives P, $\tau(P)$, $\tau^2(P)$, ... d'un point quelconque P obtenues par les puissances de τ forment une suite divergente.*

La transformation τ sans point invariant conserve le sens de la droite; la suite P, $\tau(P)$, $\tau^2(P)$, ... est donc monotone. Si elle avait un point limite, celui serait un point invariant de τ.

3. 7. *A tout couple de points A, B correspond une transformation τ de G sans point invariant qui transforme A en B.*

Soit C le milieu du segment AB. La transformation $\tau = \sigma_A \sigma_C$ change A en B et n'admet pas de point invariant (3. 5).

Nous entendons par *chaîne* d'origine $O = A_0$, engendrée par le point A_1 la suite de points

$$A_1, A_2, \ldots$$

telle que, pour tout ν, A_ν soit le milieu du segment $A_{\nu-1}A_{\nu+1}$.

Soit $A_{-1} = \sigma_O(A_1)$ le symétrique de A_1 par rapport à O, et

$$A_{-1}, A_{-2}, \ldots$$

la chaîne d'origine O engendrée par le point A_{-1}.

3. 8. *La symétrie σ_O par rapport à O change le point A_n en A_{-n}.*

A_1 est le milieu du segment OA_2; d'après la proposition 3. 3 $\sigma_O(A_1) = A_{-1}$ est le milieu du segment limité par les points $\sigma_O(O) = O$ et $\sigma_O(A_2)$; il résulte de là que $\sigma_O(A_2) = A_{-2}$. Par une induction de n à $n+1$ on déduit que $\sigma_O(A_n) = A_{-n}$, pour $n = 1, 2, 3, \ldots$

Le même raisonnement fournit le résultat suivant:

3. 9. *La symétrie par rapport au point A_k change le point A_n en A_{2k-n}.*

3. 10. *Le produit de la symétrie par rapport à O et de la symétrie par rapport à A_k transforme donc le point A_n en A_{n+2k}.*

Entre les chaînes d'origine O et les transformations du groupe G_O admettant le point invariant O, la relation suivante est valable:

3. 11. *Si la transformation T du groupe G_O change le point A_1 en A'_1, elle change le n-ième élément de la chaîne d'origine O, engendrée par A_1, en le n-ième élément de la chaîne d'origine O, engendrée par A'_1 ($n = 1, 2, 3, \ldots$).*

Soient $A_1, A_2, \ldots$ et $A'_1, A'_2, \ldots$ les éléments des deux chaînes en question. D'après la proposition 3. 3, le point $A'_1 = T(A_1)$ est le milieu du segment limité par les points $O = T(O)$ et $T(A_2)$; A'_1 est aussi le milieu du segment OA'_2; il en résulte que $A'_2 = T(A_2)$. Par une induction de n à $n+1$, on déduit similairement que $A'_n = T(A_n)$, pour tout n.

Nous introduisons sur la droite une *coordonnée* x de la façon suivante. Soient O et A_1 deux points quelconques, A_{-1} le symétrique de A_1 par rapport à O, et soient $A_1, A_2, \ldots$ et $A_{-1}, A_{-2}, \ldots$ les chaînes d'origine O, engendrées par les points A_1 et A_{-1}, respectivement. Nous attribuons à O la coordonnée $x=0$, et à A_n la coordonnée $x=n$ $(n=\pm 1, \pm 2, \ldots)$.

Désignons, pour tout entier positif n, par T_n la transformation du groupe G_O qui transforme le point A_1 en A_n. Nous désignons par $A_{1/n}$ l'image de A_1 obtenue par l'inverse de la transformation T_n, et par $A_{m/n}$ l'image de $A_{1/n}$ obtenue par la transformation T_m. Nous attribuons à $A_{m/n}$ la coordonnéc $x=m/n$, et à son symétrique par rapport à O la coordonnée $x=-m/n$.

3. 12. *La suite des points $A_{1/2}, A_{1/3}, A_{1/4}, \ldots$ converge vers O.*

Dans le cas contraire, soit Q un point, et $n_1, n_2, \ldots$ une suite d'entiers positifs tels que Q se trouve entre O et A_{1/n_ν}, pour tout $\nu: O<Q<A_{1/n_\nu}$. Il résulte de là: $O<T_{n_\nu}(Q)<T_{n_\nu}(A_{1/n_\nu})=A_1$. Soit T la transformation du groupe G_O qui change Q en A_1; désignons par A_1' l'image de A_1 obtenue par T. Considérant que les transformations T et T_{n_ν} du groupe G_O sont échangeables (3. 4), il résulte: $T_{n_\nu}T(Q)=TT_{n_\nu}(Q)=T_{n_\nu}(A_1)=A_{n_\nu}$. De la relation ci-dessus $O<T_{n_\nu}(Q)<A_1$ découle donc $O<T_{n_\nu}T(Q)<T(A_1)=A_1'$, d'où $O<A_{n_\nu}<A_1'$, pour tout ν. C'est en contradiction avec la proposition 3. 6.

Désignons par $\tau_{1/n}$ le produit de la symétrie par rapport à O et de la symétrie par rapport au point $A_{1/2n}$. Il résulte de la proposition 3. 10 que la transformation $\tau_{1/n}$ change le point $A_{m/n}$ en $A_{(m+1)/n}$. Comme la transformation $\tau_{1/n}$ conserve l'ordre des points sur la droite, nous concluons que l'ordre des points $A_{m/n}$ sur la droite est le même que celui des nombres rationnels m/n leur correspondant.

D'après 3. 10, la transformation τ_1 change le point A_0 en A_1; comme la suite $A_{1/n}$ converge vers O (3. 12), il résulte de la continuité de la transformation τ_1 que la suite $A_{(n+1)/n}$ converge vers A_1. La transformation $\tau_{1/n}$ change le point O en $A_{1/n}$ et le point A_1 en $A_{(n+1)/n}$; en conséquence de la continuité du groupe G, la transformation $\tau_{1/n}$ diffère de l'identité d'aussi peu que l'on veut, si n est suffisamment grand. Comme $\tau_{1/n}$ transforme le point $A_{m/n}$ en $A_{(m+1)/n}$, il résulte que la distance de deux points consécutifs quelconques de la suite

$$\ldots, A_{-2/n}, A_{-1/n}, O, A_{1/n}, A_{2/n}, \ldots, A_1, \ldots$$

est inférieure à un nombre positif arbitrairement petit, si n est assez grand.

De nos résultats découle immédiatement que *l'ensemble des points* $A_{m/n}$ *est partout dense sur la droite.* Nous pouvons donc étendre la coordonnée x continuement à tous les points de la droite.

Nous entendons par *translation* τ le produit de la symétrie σ_O et d'une autre symétrie quelconque. Désignons, en particulier, par $\tau_{m/n}$ le produit de la symétrie σ_O et de la symétrie par rapport au point $A_{m/2n}$. De la proposition 3. 10 nous obtenons que la translation $\tau_{m/n}$ change le point de coordonnée m'/n' en le point de coordonnée $m'/n'+m/n$. Par continuité, nous obtenons l'expression suivante de la translation $\tau_{m/n}$: $x'=x+m/n$. Cette formule s'étend par continuité à toutes les translations. Nous avons ainsi obtenu la proposition suivante:

3. 13. *Les translations contenues dans le groupe G s'expriment avec la coordonnée x par la formule:* $x'=x+b$.

En ce qui concerne l'expression des transformations du groupe G_O avec la même coordonnée, la proposition suivante est valable:

3. 14. *Les transformations du groupe* G_O *s'expriment avec la coordonnée x par la formule:* $x'=ax$.

Soient d'abord $r=m/n$ et $r'=m'/n'$ deux nombres rationnels quelconques; appliquons la proposition 3. 11 aux chaînes $A_{\nu/n'}$ et $A_{\nu m/n'n}$ $(\nu=1, 2, \ldots)$; il en résulte que la transformation de G_O qui transforme A_1 en A_r change le point $A_{r'}$ en $A_{r'r}$. En tenant compte de la continuité du groupe G_O et de celle de la coordonnée x, nous aboutissons de là au résultat que la transformation de G_O qui change A_1 en le point de coordonnée a, change le point de coordonnée x en le point de coordonnée ax.

3. 15. *Toute transformation du groupe G est le produit d'une transformation du groupe* G_O *et d'une translation.*

Soit P un point différent de O, et soit T une transformation quelconque du groupe G. Désignons par A et B les images de O et de P obtenues par T. En vertu de 3. 7 il y a une translation τ dans G qui transforme O en A. Soit ϱ la transformation de G_O qui change P en le point $\tau^{-1}(B)$. Chacune des transformations T et $\varrho\tau$ transforme O en A, et P en B; il en résulte que $T=\varrho\tau$.

L'expression de ϱ est $x'=ax$ (3. 14) et celle de $\tau: x'=x+b$

(3. 13). La transformation T s'exprime donc par la formule: $x' = ax + b$.

Le groupe G de même que le groupe des similitudes: $x' = ax + b$ sont doublement transitifs sur la droite; toute transformation $x' = ax + b$ $(a \neq 0)$ appartient, par suite, au groupe G. En résumé, *les transformations du groupe G s'expriment par la formule $x' = ax + b$ où a et b sont des nombres réels quelconques, $a \neq 0$.* C'est le théorème énoncé au début de ce paragraphe.

§ 4. Les groupes triplement transitifs de la droite projective.

Par le raisonnement du § 2 on obtient le théorème suivant:

Théorème. *Tout groupe simplement transitif de transformations topologiques de la droite projective en elle-même est homéomorphe au groupe des rotations d'une circonférence.*

Nous allons démontrer le théorème suivant:

Théorème. *Tout groupe triplement transitif de transformations topologiques de la droite projective en elle-même est homéomorphe au groupe homographique d'une variable réelle.*

Soit U un point quelconque de la droite projective; nous l'appellerons *point à l'infini*. Désignons par G_U le sous-groupe du groupe donné G formé par les transformations qui laissent invariant le point U. Le groupe G_U est *doublement transitif sur la droite* obtenue en omettant le point à l'infini de la droite projective.

A tout point O (différent de U) correspond une transformation involutive de G_U et une seule qui laisse le point O invariant (3. 2). Le groupe G contient donc une transformation involutive et une seule qui admet les points invariants O et U; nous l'appelons *involution* σ_{OU}.

Si A, B est un couple quelconque de points, il y a au moins une transformation T de G qui change O en A et U en B. La transformée $T^{-1}\sigma_{CU}T$ de σ_{OU} par T est involutive et laisse invariants les points A et B; nous appelons cette transformation *involution* σ_{AB}.

Les involutions sont liées entre elles par la *réciprocité* suivante:

4. 1. *Si l'involution σ_{AB} échange entre eux les points C et D, l'involution σ_{CD} échange entre eux les points A et B.*

Soit ϱ la transformation du groupe G qui change le triple de points (A, B, C) en (C, D, A). Comme le couple de points A, B sépare les points C et D, se correspondant par l'involution σ_{AB}, les points B et D appartiennent à un même segment AC de la droite projective. La transformation ϱ, qui échange entre eux les points A et C et transforme B en D, change, par conséquent, chacun des deux segments déterminés par les points A et C en lui-même et admet sur chacun de ces segments un point invariant, P et Q. Le carré de ϱ admet quatre points invariants: P, Q et A, C, d'où $\varrho^2 = I$.

La transformation involutive ϱ échange entre eux les points A et C, d'une part, et les points B et D, d'autre part. La transformée $\varrho^{-1}\sigma_{AB}\varrho$ de l'involution σ_{AB} par ϱ est involutive, ses points invariants sont C et D. Il résulte de là

$$\varrho^{-1}\sigma_{AB}\varrho = \sigma_{CD}.$$

L'expression à gauche montre que l'involution σ_{CD} échange entre eux les points A et B.

4. 2. *Toute transformation du groupe G qui n'appartient pas au sous-groupe G_U est le produit d'un élément T de G_U et d'une involution σ_{AB}.*

Soit S une transformation de G n'appartenant pas à G_U. Désignons par C l'image du point U obtenue par S, et soient A et B deux points qui se correspondent par l'involution σ_{CU}. Désignons par A' et B' les images de A et de B obtenues par l'inverse de la transformation S; les points A, B, A', B' sont différents de U. Si T est la transformation de G_U qui change A' en A et B' en B, et σ_{AB} l'involution ayant les points invariants A et B, qui, d'après 4. 1, échange entre eux les points C et U, la transformation $T.\sigma_{AB}$ transforme le triple de points (A', B', U) en (A, B, C); par suite $T.\sigma_{AB}$ est identique à S.

Introduisons sur la droite projective une *coordonnée* x telle qu'au point U correspond $x=\infty$, et que les transformations du groupe G_U s'expriment avec cette coordonnée par la formule $x' = ax + b$ (voir § 3).

4. 3. *Toute involution contenue dans le groupe G s'exprime par une transformation linéaire de la coordonnée x.*[3])

[3]) Concernant la méthode suivante cf. B. DE KERÉKJÁRTÓ, Sur les inversions dans un groupe commutatif, *Comptes rendus de l'Académie des Sciences, Paris,* **210** (1940), p. 288—289.

Désignons par A et A' les points de coordonnées $x = +1$ et $x = -1$. Soient P et P' les points d'affixes $x = +\xi$, et $x = -\xi$. L'involution σ_{OU} échange entre eux les points A et A' et, de même, les points P et P'. De la proposition 4. 1 il résulte que les points O et U sont échangés entre eux par chacune des deux involutions $\sigma_{AA'}$ et $\sigma_{PP'}$. La transformation $\sigma_{AA'} . \sigma_{PP'} = \mu'$ admet donc les points invariants O et U; elle appartient au sous-groupe G_{OU} formé par les transformations de G qui admettent les points invariants O et U. D'aprés 3. 14 on peut donc exprimer la transformation μ' par une équation de la forme: $x' = \xi' . x$.

Désignons par μ la transformation de G_{OU} définie par la formule: $x' = x/\xi$ qui change P en A et P' en A'. La transformée de l'involution $\sigma_{AA'}$ par μ^{-1} est l'involution $\sigma_{PP'} = \mu \sigma_{AA'} \mu^{-1}$. En multipliant cette relation à gauche par $\sigma_{AA'}$, et en tenant compte de la relation $\sigma_{AA'} . \sigma_{PP'} = \mu'$, nous obtenons la relation

$$\sigma_{AA'} . \mu \sigma_{AA'} \mu^{-1} = \mu'$$

que nous écrirons sous la forme:

$$\sigma_{AA'} \mu \sigma_{AA'} = \mu' \mu.$$

Désignons par $\sigma_{AA'}(x)$ la coordonnée du point en lequel le point d'affixe x est changé par $\sigma_{AA'}$. La transformation $\sigma_{AA'} \mu \sigma_{AA'}$ change le point d'affixe x en le point $\sigma_{AA'}\left(\sigma_{AA'}(x) . \frac{1}{\xi}\right)$. En effet $\sigma_{AA'}$ change le point x en $\sigma_{AA'}(x)$, μ change ce dernier en $\sigma_{AA'}(x) . \frac{1}{\xi}$, et $\sigma_{AA'}$ le change en $\sigma_{AA'}\left(\sigma_{AA'}(x) . \frac{1}{\xi}\right)$. D'autre part, la transformation $\mu' \mu$ change le point x en $\frac{\xi'}{\xi} . x$. De la relation ci-dessus nous obtenons donc l'égalité suivante:

$$\sigma_{AA'}\left(\sigma_{AA'}(x) . \frac{1}{\xi}\right) = \frac{\xi'}{\xi} \; x,$$

valable pour toutes valeurs réelles x et ξ. Pour déterminer ξ' en fonction de ξ, posons dans cette équation $x = 1$; comme $\sigma_{AA'}(1) = 1$, il résulte que

$$\sigma_{AA'}\left(\frac{1}{\xi}\right) = \frac{\xi'}{\xi} .$$

Posons dans l'équation ci-dessus $\xi = 1/x$, conséquemment:

$$\frac{1}{\xi} = x \quad \text{et} \quad \frac{\xi'}{\xi} = \sigma_{AA'}(x) ;$$

nous obtenons l'équation suivante :

$$\sigma_{AA'}(\sigma_{AA'}(x) \,.\, x) = \sigma_{AA'}(x) \,.\, x.$$

Cette équation signifie que, pour toute valeur de x, le point de coordonnée $\sigma_{AA'}(x) \,.\, x$ est un point invariant de $\sigma_{AA'}$, il coïncide donc ou avec le point $+1$ ou avec -1 c'est-à-dire que $\sigma_{AA'}(x) = \pm 1/x$. Pour $x = 1$, on a $\sigma_{AA'}(x) = +1/x$, et comme $\sigma_{AA'}(x)$ est continue, on obtient, pour toute valeur de x, l'expression suivante :

$$\sigma_{AA'}(x) = +\frac{1}{x}.$$

Soit B, C un autre couple quelconque de points (différents de U). Soit T la transformation du groupe G_U qui change A en B et A' en C. En conséquence de la formule $\sigma_{BC} = T^{-1}\sigma_{AA'}T$, l'involution σ_{BC} est le produit des transformations linéaires T^{-1}, $\sigma_{AA'}$ et T, par suite, σ_{BC} s'exprime aussi par une transformation linéaire.

En tenant compte de la proposition 4. 2 nous concluons que toute transformation du groupe G s'exprime par une transformation linéaire de la coordonnée x. Le groupe G ainsi que le groupe des transformations linéaires $x' = (ax+b)/(cx+d)$ sont triplement transitifs sur la droite projective ; toute transformation linéaire appartient donc au groupe G.

Nous avons ainsi démontré que *les transformations du groupe G s'expriment par les transformations linéaires $x' = (ax+b)/(cx+d)$ où a, b, c, d sont des nombres réels quelconques tels que $ad - bc \neq 0$.* C'est le théorème énoncé au début de ce paragraphe.

(Reçu le 24 juin 1940)

Über die reellen Nullstellen gewisser Polynome mit Parametern.

Von GYULA (JULIUS) V. SZ. NAGY in Kolozsvár.

1. Einleitung.

Bei der Darstellung algebraischer Kurven vom Maximalindex in mehrdimensionalen Räumen habe ich einige algebraische Sätze gefunden, die mir als neu und nicht ohne gewissem Interesse zu sein scheinen.

Der Grund dieser Sätze liegt im folgenden leicht beweisbaren Satze:

I. *Bedeuten $f(x)$, $g(x)$, $F(x)$ und $G(x)$ Polynome mit reellen Koeffizienten und sind a_h und a_i solche reelle Nullstellen von $f(x)$, zwischen denen eine ungerade bzw. gerade Anzahl der Nullstellen von $g(x)$ bzw. $F(x)$ liegt, so hat das Polynom*

$$\varphi(x) = g(x)\,F(x) - f(x)\,G(x)$$

im abgeschlossenen Intervall (a_h, a_i) mindestens eine Nullstelle.

Aus diesem Satze läßt sich (unter anderem) auch der Satz ableiten:

II. *Sind $f_m(x)$, $g_{m-1}(x)$, $F_k(x)$ und $G_{k-1}(x)$ Polynome mit reellen Koeffizienten und vom Grade m, $m-1$, k bzw. $k-1$ $(k<m)$ und hat das Polynom $f_m(x)$ nur reelle und verschiedene Nullstellen, die von den ebenfalls reellen Nullstellen des Polynoms $g_{m-1}(x)$ getrennt werden, so hat das Polynom*

$$\varphi(x) = g_{m-1}(x)\,F_k(x) - f_m(x)\,G_{k-1}(x)$$

zwischen der kleinsten und größten Nullstelle von $f_m(x)$ mindestens $m-k-1$ Nullstellen. Das Polynom $\varphi(x)$ hat also für beliebige Polynome $F_k(x)$ und $G_{k-1}(x)$ mit reellen Koeffizienten höchstens $m+k-1$ und mindestens $m-k-1$ reelle Nullstellen (vorausgesetzt, daß nicht jeder Koeffizient von $F_k(x)$ und $G_{k-1}(x)$ verschwindet).

Aus diesem Satze folgt der Satz:

III. *Sind* $f_m(x)$, $F_k(x)$ *und* $G_{k-1}(x)$ *Polynome mit reellen Koeffizienten und vom Grade* m, k *bzw.* $k-1$ $(k<m)$ *und hat das Polynom* $f_m(x)$ *nur reelle Nullstellen, so hat das Polynom*

$$\varphi(x)=f'_m(x)\,F_k(x)-f_m(x)\,G_{k-1}(x)$$

für beliebige $F_k(x)$ *und* $G_{k-1}(x)$ *mit reellen Koeffizienten mindestens* $m-k-1$ *und höchstens* $m+k-1$ *reelle Nullstellen.*

Die für die Minimalanzahl der reellen Nullstellen ausgesprochenen Sätze lassen sich ohne Weiteres auf stetige Funktionen von entsprechenden Eigenschaften verallgemeinern.

2. Beweis der Sätze I, II und III.

Nach den Annahmen des Satzes I bestehen die Ungleichungen

(1) $$g(a_h)\,g(a_i)\leqq 0 \quad \text{und} \quad F(a_h)\,F(a_i)\geqq 0,$$

weil das Polynom $g(x)$ bzw. $F(x)$ im Innern des Intervalles (a_h, a_i) eine ungerade bzw. gerade Anzahl von Nullstellen hat.

Ist nun

$$g(a_h)\,g(a_i)\,F(a_h)\,F(a_i)\neq 0,$$

so haben

$$\varphi(a_h)=g(a_h)\,F(a_h) \quad \text{und} \quad \varphi(a_i)=g(a_i)\,F(a_i)$$

nach (1) entgegengesetzte Vorzeichen. Dann hat also $\varphi(x)$ im Innern des Intervalles (a_h, a_i) eine ungerade Anzahl von Nullstellen, also mindestens eine Nullstelle.

Ist aber

$$g(a_h)\,F(a_h)\,g(a_i)\,F(a_i)=\varphi(a_h)\,\varphi(a_i)=0,$$

so ist mindestens ein Endpunkt des Intervalles (a_h, a_i) eine Nullstelle des Polynoms $\varphi(x)$.

Damit ist der Satz I vollständig bewiesen.

Aus diesem Satz folgt auch der Satz:

IV. *Bedeuten* $f(x)$, $g(x)$, $F(x)$ *und* $G(x)$ *Polynome mit reellen Koeffizienten, sind ferner* a_h *und* a_i *solche reelle Nullstellen von* $f(x)$, *zwischen denen eine ungerade Anzahl der Nullstellen von* $g(x)$ *liegt und hat endlich das Polynom* $\varphi(x)=g(x)\,F(x)-f(x)\,G(x)$ *im abgeschlossenen Intervall* (a_h, a_i) *keine Nullstelle, so hat das*

Polynom $F(x)$ im Intervall (a_h, a_i) eine ungerade Anzahl von Nullstellen, also mindestens eine Nullstelle.

Für den Beweis des Satzes II bezeichnen wir mit $a_1, a_2, \ldots, a_m$ die der Größe nach geordneten Nullstellen des Polynoms $f_m(x)$.

Verschwindet das Polynom $F_k(x)$ in keinem der Punkte $a_1, a_2, \ldots, a_m$, so können höchstens k der $m-1$ Intervalle (a_1, a_2), $(a_2, a_3), \ldots, (a_{m-1}, a_m)$ eine ungerade Anzahl der Nullstellen von $F_k(x)$ enthalten. Es gibt also unter den $m-1$ Intervallen mindestens $m-k-1$, die eine gerade Anzahl der Nullstellen von $F_k(x)$ enthalten. Jedes dieser Intervalle enthält nach Satz I mindestens eine Nullstelle des Polynoms $\varphi(x)$, weil die Polynome $g_{m-1}(x)$ und $F_k(x)$ in den Punkten $a_1, a_2, \ldots, a_m$ nicht verschwinden. Im Intervall (a_1, a_m) hat also das Polynom $\varphi(x)$ mindestens $m-k-1$ Nullstellen.

Damit ist Satz II für den Fall $F(a_1) F(a_2) \ldots F(a_m) \neq 0$ bewiesen. Es bleibt nur der Fall übrig, daß $F_k(x)$ in einigen der Punkte $a_1, a_2, \ldots, a_m$ verschwindet.

Ist $F_k(a_h) = 0$, so ist der Punkt a_h offenbar eine Nullstelle des Polynoms $\varphi(x)$.

Fallen nun s Nullstellen des Polynoms $F_k(x)$ mit s Nullstellen von $f_m(x)$ zusammen, so stimmen sie auch mit s Nullstellen von $\varphi(x)$ überein. Zum Beweis des Satzes II können wir annehmen, daß $s < m-k-1$ ist. Die der Größe nach aufeinanderfolgenden Paare der von den Nullstellen des Polynoms $F_k(x)$ verschiedenen $m-s$ Nullstellen von $f_m(x)$ bestimmen $m-s-1$ Intervalle. Unter diesen Intervallen gibt es mindestens $m-2s-1$ solche, die im Innern keine Nullstelle von $f_m(x)$ haben. Unter diesen letzteren Intervallen gibt es also mindestens $(m-2s-1)-(k-s) = (m-k-1)-s$ solche, die nicht nur in den Endpunkten, sondern auch im Innern keine Nullstelle von $F_k(x)$ enthalten. (Das Polynom $F_k(x)$ hat nämlich höchstens $k-s$ solche reelle Nullstellen, die von den Nullstellen von $f_m(x)$ abweichen.) Jedes Intervall von dieser letzteren Eigenschaft enthält nach dem Satze I mindestens eine Nullstelle von $\varphi(x)$ in seinem Innern. Das Polynom $\varphi(x)$ hat also im Intervall (a_1, a_m) mindestens $m-k-1$ reelle Nullstellen, weil sie in den Punkten $a_1, a_2, \ldots, a_m$ s Nullstellen und im Innern der Intervalle $(a_1, a_2), (a_2, a_3), \ldots, (a_{m-1}, a_m)$ mindestens $m-k-1-s$ Nullstellen besitzt.

Damit ist der Satz II bewiesen.

Das Polynom $\varphi(x)$ kann höchstens $m+k-1$ Nullstellen haben, weil es offenbar nur dann identisch verschwindet, wenn beide Polynome $F_k(x)$ und $G_{k-1}(x)$ identisch verschwinden. Dies folgt daraus, daß die Polynome $f_m(x)$ und $g_{m-1}(x)$ teilerfremd sind und daß der Grad von $F_k(x)$ bzw. $G_{k-1}(x)$ kleiner als derjenige von $f_m(x)$ bzw. $g_{m-1}(x)$ ist.

Wir führen die folgende *Benennung* ein: Die Nullstellen eines Polynoms $f_m(x)$ vom m-ten Grade mit lauter reellen Nullstellen sind durch die Nullstellen eines Polynoms $g_{m-1}(x)$ vom $(m-1)$-ten Grade getrennt, falls jede i-fache Nullstelle $(i>1)$ von $f_m(x)$ eine $(i-1)$-fache Nullstelle von $g_{m-1}(x)$ ist und wenn $g_{m-1}(x)$ zwischen zwei benachbarten (verschiedenen) Nullstellen von $f_m(x)$ genau einmal verschwindet.

Mit dieser Benennung gilt die folgende Ergänzung des Satzes II:

V. *Der Satz II gilt auch dann, wenn das Polynom $f_m(x)$ auch mehrfache Nullstellen hat.*

Bezeichnet nämlich $h(x)$ den größten gemeinsamen Teiler der Polynome $f_m(x)$ und $g_{m-1}(x)$, so ist $h(x)$ auch ein Teiler von $\varphi(x)$.

Sind

$$f_m(x)=h(x)f_\mu(x),\ g_{m-1}(x)=h(x)g_{\mu-1}(x) \quad \text{und} \quad \varphi(x)=h(x)\,\psi(x),$$

so ist

$$\psi(x)=g_{\mu-1}(x)\,F_k(x)-f_\mu(x)\,G_{k-1}(x).$$

Hier ist $m=r+\mu$, wenn r den Grad von $h(x)$ bedeutet.

Ist $k\leqq\mu-1$, so hat $\psi(x)$ zwischen der kleinsten und größten Nullstelle von $f_\mu(x)$ (nach Satz II) mindestens $\mu-k-1$ Nullstellen. Das Polynom $\varphi(x)$ hat also im abgeschlossenen Intervall zwischen der kleinsten und der größten Nullstelle von $f_\mu(x)$ mindestens $(\mu-k-1)+r=m-k-1$ Nullstellen. Das gleiche gilt auch im Falle $k>\mu-1$, weil dann $r>r+(\mu-k-1)=m-k-1$ ist.

Damit ist der Satz V vollständig bewiesen. Aus diesem Satz folgt der Satz III unmittelbar.

3. Weitere Sätze.

Aus dem Beweise der Sätze I und II folgen auch die Sätze:

VI. *Sind $f_m(x)$, $g_{m-1}(x)$, $F_k(x)$ und $G_{k-1}(x)$ Polynome mit reellen Koeffizienten und vom Grade m, $m-1$, k bzw. $k-1$, hat ferner das Polynom $f_m(x)$ lauter reelle und verschiedene Nullstellen, die* *von den Nullstellen des Polynoms* $g_{m-1}(x)$ *getrennt werden, und gibt es endlich p bzw. q Paare der benachbarten Nullstellen von* $f_m(x)$ *bzw.* $g_{m-1}(x)$, *deren Intervalle eine ungerade Anzahl der Nullstellen von* $F_k(x)$ *bzw.* $G_{k-1}(x)$ *enthalten, so hat das Polynom*

$$\varphi(x) = g_{m-1}(x)\, F_k(x) - f_m(x)\, G_{k-1}(x)$$

mindestens $r = \mathrm{Max}\,\{m-1-p,\ m-2-q\}$ *reelle Nullstellen.*

VII. *Sind* $f(x)$, $g(x)$, $F(x)$ *und* $G(x)$ *Polynome mit reellen Koeffizienten, haben ferner die Polynome* $f(x)$ *und* $g(x)$ *mindestens m bzw. n solche Paare der benachbarten reellen Nullstellen, deren Intervalle eine ungerade Anzahl der Nullstellen von* $g(x)$ *bzw.* $f(x)$ *enthalten, und hat endlich das Polynom* $F(x)$ *bzw.* $G(x)$ *höchstens p bzw. q reelle Nullstellen, so hat das Polynom*

$$\varphi(x) = g(x)\, F(x) - f(x)\, G(x)$$

mindestens $r = \mathrm{Max}\,\{m-p,\ n-q\}$ *reelle Nullstellen.*

VIII. *Sind* $f(x)$, $g(x)$, $F(x)$ *und* $G(x)$ *Polynome mit reellen Koeffizienten, hat ferner das Polynom* $f(x)$ *m solche Paare von benachbarten reellen Nullstellen, deren Intervalle eine ungerade Anzahl der reellen Nullstellen von* $g(x)$ *enthalten und ist endlich* $F(x)$ *ein beliebiges definites oder semidefinites Polynom, so hat jedes Polynom* $\varphi(x)$ *von der Form*

$$\varphi(x) = g(x)\, F(x) - f(x)\, G(x)$$

mindestens m reelle Nullstellen.

Ein definites Polynom hat nämlich keine reelle Nullstelle und jede reelle Nullstelle eines semidefiniten Polynoms besitzt eine gerade Multiplizität.

IX. *Sind* $f_m(x)$, $g_{m-1}(x)$, $F_k(x)$ *und* $G_{k-1}(x)$ *Polynome mit reellen Koeffizienten und vom Grade* m, $m-1$, k *bzw.* $k-1$ $(k < m)$ *und hat das Polynom* $f_m(x)$ *lauter reelle Nullstellen, die von den Nullstellen des Polynoms* $g_{m-1}(x)$ *getrennt werden, so haben die Kurven*

$$y\, g_{m-1}(x) - f_m(x) = 0 \quad \textit{und} \quad y\, G_{k-1}(x) - F_k(x) = 0$$

(außer den in den unendlichfernen Punkt der y-Achse fallenden $(m-1)(k-1)$ Punkten) *mindestens* $m-k-1$ *und höchstens* $m+k-1$ *reelle Treffpunkte.*

Die Abzissen der Treffpunkte dieser Kurven genügen nämlich der Gleichung

$$\varphi(x) = g_{m-1}(x)\, F_k(x) - f_m(x)\, G_{k-1}(x) = 0.$$

X. *Sind* $f_m(x)$, $g_{m-2}(x)$, $F_k(x)$ *und* $G_{k-1}(x)$ *Polynome mit reellen Koeffizienten und vom Grade* m, $m-2$, k *bzw.* $k-1$, *und hat das Polynom* $f_m(x)$ *lauter reelle und verschiedene Nullstellen, von denen die* $m-1$ *kleinsten von den Nullstellen des Polynoms* $g_{m-2}(x)$ *getrennt werden, so haben die Kurven*

$$y^2 g_{m-2}(x) - f_m(x) = 0 \text{ und } y G_{k-1}(x) - F_k(x) = 0$$

(außer den in den unendlichfernen Punkt der y-Achse fallenden $(m-2)(k-1)$ Punkten) *mindestens* $m-2$ *und höchstens* $m-2+2k$ *reelle Treffpunkte.*

Die Abszissen der Treffpunkte dieser Kurven genügen nämlich der Gleichung

$$\varphi(x) = g_{m-2}(x) F_k^2(x) - f_m(x) G_{k-1}^2(x) = 0,$$

die nach dem Satz VIII mindestens $m-2$ reelle Wurzeln besitzt.

(*Eingegangen am 22. August 1940.*)

Intégration de deux équations aux différences finies linéaires à deux variables.

Par DANIEL ARANY à Budapest.

Dédié à M. Louis Bachelier.

Premier Problème.

Envisageons l'équation suivante

$$(1) \qquad y(x, t) = p\, y(x-1, t+1) + q\, y(x-1, t-1)$$

aux conditions limites :

$y(x, 0) = 0,\ \ y(x, n) = 0,\ \ y(x, x) = q^x$ et $y(x, t) = 0$ si $x < t$.

Soit en outre $p + q = 1$.

Considérons la solution particulière suivante

$$(2) \qquad y(x, t) = \alpha^x \beta^t.$$

De (1) on tire

$$(3) \qquad \alpha^x \beta^t = p\, \alpha^{x-1} \beta^{t+1} + q\, \alpha^{x-1} \beta^{t-1}$$

et en simplifiant

$$(4) \qquad \beta^2 - \frac{\alpha}{p} \beta + \frac{q}{p} = 0.$$

Désignons les racines de l'équation (4) par β_1 et β_1. En multipliant cette équation successivement par $\beta, \beta^2, \ldots, \beta^{t-2}$ on obtient $t-2$ équations de la forme

$$(5) \qquad \beta^t = \beta\, u(t) + u_1(t)$$

où $u(t)$ et $u_1(t)$ sont des fonctions de α, p et q.

L'équation (5) subsiste pour β_1 et β_2. On a donc

$$\beta_1^t = \beta_1 u(t) + u_1(t) \quad \text{et} \quad \beta_2^t = \beta_2 u(t) + u_1(t)$$

il en résulte

$$u(t)=\frac{\beta_1^t-\beta_2^t}{\beta_1-\beta_2} \tag{6}$$

$$u_1(t)=-\frac{q}{p}u(t-1);$$

on a évidemment

$$u(0)=0 \quad \text{et} \quad u(1)=1.$$

De l'équation (4) on obtient

$$\beta=\frac{1}{2p}(\alpha\pm\sqrt{\alpha^2-4pq})$$

ce qui donne, en posant

$$\alpha=2\sqrt{pq}\cos\vartheta, \tag{7}$$

$$\beta_1=\left(\frac{q}{p}\right)^{\frac{1}{2}}(\cos\vartheta+i\sin\vartheta) \quad \text{et} \quad \beta_2=\left(\frac{q}{p}\right)^{\frac{1}{2}}(\cos\vartheta-i\sin\vartheta), \tag{8}$$

de plus

$$\beta_1^t=\left(\frac{q}{p}\right)^{\frac{t}{2}}(\cos t\vartheta+i\sin t\vartheta) \quad \text{et} \quad \beta_2^t=\left(\frac{q}{p}\right)^{\frac{t}{2}}(\cos t\vartheta-i\sin t\vartheta)$$

et enfin il vient de (6)

$$u(t)=\left(\frac{q}{p}\right)^{\frac{t-1}{2}}\frac{\sin t\vartheta}{\sin\vartheta}. \tag{9}$$

En multipliant l'équation (5) par α^x on trouve

$$y(x,t)=u(t)\,y(x,1)+u_1(t)\,y(x,0);$$

comme $y(x,0)=0$, on a

$$y(x,t)=u(t)\,y(x,1)$$

Comme on doit avoir en outre $y(x,n)=0$, on a en vertu de (9)

$$\frac{\sin\vartheta n}{\sin\vartheta}=0; \tag{10}$$

ϑ ne peut donc avoir que les valeurs $k\pi/n$ $(k=1,2,\ldots,n-1)$. On obtient donc la solution suivante de l'équation (1) satisfaisant à $y(x, 0)=0$ et $y(x, n)=0$:

$$y(x, t)=\sum_{k=1}^{n-1} C_k \left(\frac{q}{p}\right)^{\frac{t-1}{2}} \frac{\sin \frac{k\pi t}{n}}{\sin \frac{k\pi}{n}} y(x, 1).$$

Remplaçons $y(x, 1)$ par sa solution particulière suivante

$$y(x, 1)=\alpha^x(\beta_1+\beta_2)=\left(2\sqrt{pq}\cos\frac{k\pi}{n}\right)^x 2\cos\frac{k\pi}{n}\left(\frac{q}{p}\right)^{\frac{1}{2}}$$

ce qui donne

$$y(x, t)=\left(\frac{q}{p}\right)^{\frac{t}{2}}(qp)^{\frac{x}{2}}\sum_{k=1}^{n-1} C_k \frac{\sin \frac{k\pi t}{n}}{\sin \frac{k\pi}{n}}\left(2\cos\frac{k\pi}{n}\right)^{x+1}.$$

Pour $x=1$ et $t=1$ on doit avoir $y(1, 1)=q$. Il en résulte

$$\frac{1}{4}=\sum_{k=1}^{n-1} C_k \cos^2\frac{k\pi}{n}.$$

De plus pour $x=1$ et $t=2, 3, 4, \ldots, n-1$ on doit avoir $y(1, t)=0$, par suite

$$0=\sum_{k=1}^{n-1} C_k \cos^2\frac{k\pi}{n} \frac{\sin\frac{k\pi t}{n}}{\sin\frac{k\pi}{n}}.$$

Posons

$$A_k=C_k\frac{\cos^2\frac{k\pi}{n}}{\sin\frac{k\pi}{n}}.$$

Alors les équations à satisfaire seront:

$$\sum_{k=1}^{n-1} A_k \sin\frac{k\pi}{n}=\frac{1}{4},$$

$$\sum_{k=1}^{n-1} A_k \sin\frac{\pi k t}{n}=0$$

pour $t=2, 3, \ldots, n-1$. On démontre au Calcul des Différences Finies que ces équations sont satisfaites si

$$A_k=\frac{1}{2n}\sin\frac{k\pi}{n},$$

pourvu que l'on ait $n>2$.[1]) On en conclut que

[1]) Voir p. e. Ch. Jordan, *Calculus of Finite Differences* (Budapest, 1939), p. 127.

$$C_k = \frac{1}{2n} \frac{\sin^2 \frac{k\pi}{n}}{\cos^2 \frac{k\pi}{n}}$$

et finalement on a

$$(11) \qquad y(x, t) = \frac{2^x}{n} q^{\frac{t+x}{2}} p^{\frac{x-t}{2}} \sum_{k=1}^{n-1} \sin \frac{k\pi}{n} \sin \frac{k\pi t}{n} \left(\cos \frac{k\pi}{n}\right)^{x-1}.$$

Comme cette valeur a été obtenue à l'aide de

$$y(1, 1) = q \quad \text{et} \quad y(1, t) = 0 \quad \text{si} \quad t > 1$$

il reste encore à vérifier que l'on a aussi

$$y(x, x) = q^x \quad \text{et} \quad y(x, t) = 0 \quad \text{si} \quad t > x.$$

On a en exprimant les puissances des cosinus par les cosinus des multiples de l'angle

$$2^{x-2} (\cos \vartheta)^{x-1} = \sum_{j=0} \binom{x-1}{j} \cos (x - 1 - 2j) \vartheta,$$

puis on a

$$\cos(x - 1 - 2j) \vartheta \,.\, \sin \vartheta = \frac{1}{2} [\sin(x - 2j) \vartheta + \sin (x - 2j - 2) \vartheta].$$

Il en résulte, que dans l'expression (11) le coefficient le plus élévé de ϑ sera x; de plus pour $t = x$ la somme de tous les autres termes sera nulle à cause de

$$\sum_{k=1}^{n-1} \sin \frac{xk\pi}{n} \sin \frac{\nu k\pi}{n} = 0 \quad \text{si} \quad x \neq \nu$$

et l'on aura

$$y(x, x) = \frac{2}{n} q^x \sum_{k=1}^{n-1} \sin^2 \frac{xk\pi}{n} = q^x.$$

Lorsque $t > x$, tous les termes seront nuls pour la même raison[2]).

Second Problème.

Cherchons l'intégrale de l'équation

$$z(x, t) = p\, z(x - 1, t + 1) + q\, z(x - 1, t - 1)$$

identique à (1); mais aux conditions limites suivantes:

$$z(x, 0) = 0, \; z(x, n) = 0, \; z(1, \alpha - 1) = p, \; z(1, \alpha + 1) = q$$
$$(1 < \alpha < n - 1).$$

---

[2]) La démonstration a été considérablement abrégée grâce à une remarque de M. Charles Jordan.

Dans ce cas on a

$$(12) \quad z(x, t) = \frac{2^{x+1}}{n} p^{\frac{1}{2}(x-t+\alpha)} q^{\frac{1}{2}(x+t-\alpha)} \sum_{\lambda=1}^{n-1} \sin\left(\alpha \frac{\lambda\pi}{n}\right) \sin\left(t \frac{\lambda\pi}{n}\right) \cos^x\left(\frac{\lambda\pi}{n}\right).$$

En comparant ce résultat à (11) on trouve que

$$y(x, t) = q[z(x-1, t)]_{\alpha=1}.$$

Applications des deux résultats trouvés.

La quantité $y(x, t)$ représente la probabilité de la perte de t francs en x parties, d'un joueur A, possédant t francs, qui joue contre un joueur B, possédant $n-t$ francs, p étant la chance de gagner du joueur A à chaque partie et q celle du joueur B.

La quantité $z(x, t)$ représente la probabilité de la perte de $t-\alpha$ francs en x parties, d'un joueur A, possédant t francs au commencement du jeu, qui joue contre un joueur B, possédant $n-t$ francs au commencement du jeu. Les chances de gagner à chaque partie sont pour les deux joueurs les mêmes que dans le cas précédent.

Robert L. Ellis, qui a établi le premier la formule (12) dans un Mémoire intitulé: "On the solution of equations of finite differences" (*The Mathematical and other writings of R. Leslie Ellis,* Cambridge and London, 1863) remarque au p. 210: "One point, which is worth perhaps notice, is the *symmetrical* manner in which α and t, and p and q enter into $z(x, t)$; the result, however, which is the interpretation of this symmetry may probably be obtained by general consideration."

Dans un mémoire[3]) paru en 1933 j'ai traité "Le problème des parcours" que j'ai formulé ainsi: "Le problème des parcours consiste dans la recherche de la probabilité, d'arriver du point 0 à un point x en n pas, en touchant le point x *au moins une fois* et ne touchant *jamais* certains points que j'ai nommé points extrémaux". (La probabilité de passer d'un point à *l'un* des points voisins soit $p = q = \frac{1}{2}$.)

[3]) D. Arany, Le problème des parcours, *Tôhoku Math. Journ.* 37 (1933), p. 17—22.

Dans un second Mémoire[4]) paru la même année (1933) j'ai établi une formule qui définit la probabilité de passer d'un point x au point y en n pas, sous les mêmes conditions qui sont valables pour le passage de 0 à x. Si j'applique cette formule, qui coïncide (aux notations près) avec la formule (12) du mémoire précédent, la symétrie de α et t est expliquée, puisque la probabilité de passer α à t, ou celle de passer de t à α est la même.

Budapest, le 27 Janvier 1939.

(Reçu le 4 novembre 1939)

---

[4]) D. ARANY, Le problème de parcours, *Association Française pour l'Avancement des Sciences, Compte Rendu de le* 57e *Session* (Chambéry, 1933), p. 20—23.

Über Funktionen, die ein endliches Dirichletsches Integral haben.

Von A. Sólyi in Szeged.

Einleitung.

Es sei in dem n-dimensionalen Raum eine quadratisch integrierbare Funktion $g(x_1, x_2, \ldots, x_n)$ gegeben. Die Integrale

$$\int\limits_T \left(\frac{g(x_1, x_2, \ldots, x_\alpha + h, \ldots, x_n) - g(x_1, \ldots, x_n)}{h}\right)^2 dV$$

(T bedeutet einen n-dimensionalen Bereich, $dV = dx_1 dx_2 \ldots dx_n$) seien gleichmäßig beschränkt für alle h $(\alpha = 1, 2, \ldots, n)$.

Wegen variationstheoretischer Anwendungen bieten die folgenden Fragen ein gewisses Interesse: Was kann man von den partiellen Ableitungen von g sagen? Für welche $\nu > 2$ kann man folgern, daß $\int\limits_T |g|^\nu dV$ endlich bleibt? (Diese Exponenten werden wir „Integrabilitätsexponenten" nennen.)

Wir werden sehen, daß unsere Annahme äquivalent damit ist, daß die ersten partiellen Ableitungen der Funktion g fast überall existieren und quadratisch integrierbar sind. Oder, m. a. W., daß ihr, in geeigneter Weise definiertes Dirichletsches Integral endlich ist. Es wird sich auch ergeben, daß die Differenzenquotienten im quadratischen Mittel gegen die entsprechenden Ableitungen konvergieren und daß g fast überall gleich dem Lebesgueschen Integral jeder seiner Ableitungen ist.

Was die Integrabilitätsexponenten betrifft, so wird bewiesen, daß alle ν Integrabilitätsexponenten sind, die der Ungleichung $\nu < \frac{2n}{n-2}$ genügen. Für $\nu > \frac{2n}{n-2}$ können, wie es ein einfaches

Beispiel zeigt, die Integrale $\int_T |g|^\nu \, dV$ divergent werden. Den Fall $\nu = \frac{2n}{n-2}$ konnte ich nicht entscheiden.

Zur Behandlung dieser Fragen habe ich auf Anregung des Herrn Prof. F. RIESZ g in eine mehrfache Fouriersche Reihe entwickelt und den Beweis mit gewißen Abschätzungen der Fourierschen Koeffizienten durchgeführt. Wir mußten uns daher zuerst auf den Fall beschränken, wo g in allen Veränderlichen periodisch mit der Periode 2π ist und der Bereich T mit einem Periodenwürfel W von g identisch ist. Der allgemeine Fall wird mit Hilfe einer geeigneten Fortsetzung auf diesen einfacheren Fall zurückgeführt.

1. Über die Ableitungen von g.

Bezeichnungen.

$$D_\alpha^h g = \frac{g(x_1, x_2, \ldots, x_\alpha + h, \ldots, x_n) - g(x_1, \ldots, x_n)}{h}.$$

Es sei S die Menge derjenigen Punkte, die außerhalb T und von T in einer Entfernung kleiner als d liegen; d ist eine willkürlich gewählte positive Konstante.

Voraussetzung. Alle Integrale

$$\int_{T+S} (D_\alpha^h g)^2 \, dV$$

sind gleichmäßig beschränkt.

Satz. *Die partiellen Ableitungen von g existieren in T fast überall. Sie sind alle quadratisch integrierbar. Die Differenzenquotienten von g streben im quadratischen Mittel gegen die entsprechenden Ableitungen, d. i.*

$$\int_T (D_\alpha^h g - g_{x_\alpha})^2 \, dV \to 0, \quad \textit{wenn} \quad h \to 0.$$

Beweis für den periodischen Fall ($T = W$). g wird in eine mehrfache Fouriersche Reihe entwickelt:

$$g \sim \sum_{l_1, \ldots, l_n = -\infty}^{\infty} a_{l_1 \ldots l_n} e^{i(l_1 x_1 + \ldots + l_n x_n)}.$$

Wir führen der Kürze halber die folgenden Bezeichnungen ein:

$$a_{l_1 \ldots l_n} = a_l,$$
$$e^{i(l_1 x_1 + \ldots + l_n x_n)} = e^{i l x}.$$

Es genügt aus Symmetriegründen nur die nach x_1 genommenen Differenzenquotienten zu untersuchen. Es ist

$$D_1^h g \sim \sum_{(l)} a_l \frac{e^{il_1 h}-1}{h} e^{ilx}.$$

Durch Anwendung des Parsevalschen Satzes bekommt man:

$$\frac{1}{(2\pi)^n}\int_W (D_1^h g)^2 dV = \sum_{(l)} |a_l|^2 \frac{(e^{il_1 h}-1)(e^{-il_1 h}-1)}{h^2}.$$

Da nach unserer Annahme unabhängig von h

$$\int_W (D_1^h g)^2 dV \leqq M$$

gilt, so folgt, daß die Reihe

$$\sum_{(l)} |a_l|^2 l_1^2$$

konvergent ist.

Es gibt daher auf Grund des Riesz—Fischerschen Satzes eine Funktion g_1, die die Fourierkoeffizienten $l_1 a_l$ besitzt. Wir wollen zeigen, daß in T fast überall

$$g_1 = \frac{\partial g}{\partial x_1}.$$

Wir schätzen dazu das Quadratintegral von $g_1 - D_1^h g$ ab. Wegen

$$g_1 - D_1^h g \sim \sum_{(l)} a_l \left(\frac{e^{il_1 h}-1}{h} - l_1\right) e^{ilx}$$

hat man

$$\frac{1}{(2\pi)^n}\int_W (g_1 - D_1^h g)^2 dV = \sum_{(l)} |a_l|^2 \left|\frac{e^{il_1 h}-1}{h} - l_1\right|^2.$$

Wir spalten die Summe in zwei Teile, je nachdem $|l_1| \leqq N$ oder $> N$ ist.

Der erste Teil konvergiert offenbar gegen 0, wenn $h \to 0$. Die zweite Summe ist kleiner, als

$$\sum_{|l_1| > N} 4|a_l|^2 l_1^2.$$

Wenn wir N ins Unendliche wachsen lassen, dann strebt auch diese Summe wegen der Konvergenz der Reihe

$$\sum_{l_1,\ldots,l_n=-\infty}^{+\infty} |a_l|^2 l_1^2$$

gegen 0. Es folgt daher:

$$\int_W (g_1 - D_1^h g)^2 \, dV \to 0,$$

wenn $h \to 0$. Aus dieser Relation folgt aber die Gültigkeit unseres Satzes.

2. Über die Integrierbarkeitsexponenten von g.

Es folgt für $n = 1$ unmittelbar aus dem vorigen Satze, daß g das Integral seiner Ableitung und somit stetig ist. Man sieht aus dem Beispiel $g = \log^{0 \cdot 1} r$ mit $r = \sqrt{x_1^2 + \ldots + x_n^2}$, daß die Stetigkeit von g schon für $n = 2$ nicht gefolgert werden kann. Es folgt aber, daß die Integrierbarkeitsexponenten von g beliebig groß sind. Für $n > 2$ liegen die Integrierbarkeitsexponenten unter einer endlichen Schranke, welche gegen 2 strebt, wenn n ins Unendliche wächst. Es gilt nämlich der

Satz. *Aus den gemachten Voraussetzungen über die Differenzenquotienten folgt, daß $\int_T |g|^\nu \, dV$ endlich ist für jedes positive ν, welche der Ungleichung*

$$\nu < \frac{2n}{n-2}$$

genügt ($n \neq 1, 2$). *Für $n = 1$ ist g stetig, für $n = 2$ kann ν beliebig groß gewählt werden.*

Bemerkung. Man kann die in dem Satz vorkommende obere Schranke nicht verkleinern. Das sieht man aus dem Gegenbeispiel:

$$g = \frac{1}{r^{\frac{n-2}{2} - \varepsilon}}.$$

Diese Funktion genügt unseren Voraussetzungen für jedes positive ε, dennoch kann ε für jedes $\nu > \dfrac{2n}{n-2}$ so klein gewählt werden, daß $\int |g|^\nu \, dV$ divergiere. Es wäre interessant und wegen der variationstheoretischen Anwendungen wichtig, auch den Fall $\nu = \dfrac{2n}{n-2}$ zu untersuchen. Es ist mir aber die Entscheidung dieses Falles nicht gelungen.

B e w e i s für den periodischen Fall ($T = W$). Es folgt aus dem Young—Hausdorffschen Satze, daß, wenn zwei positive Zahlen, μ und ν, der Relation

$$\frac{1}{\mu} + \frac{1}{\nu} = 1$$

genügen, $\mu \leqq \nu$, und die Reihe $\sum_{(l)} |a_l|^\mu$ konvergiert, dann konvergiert auch das Integral

$$\int_W |g|^\nu \, dV.$$

Wir haben also, um unseren Satz zu beweisen, nur die Konvergenz der Reihe

$$\sum_{l_1, \ldots, l_n = -\infty}^{\infty} |a_l|^\mu$$

für jedes

$$\mu > \frac{1}{1 - \frac{n-2}{2n}} = \frac{2n}{n+2}$$

zu zeigen.

Da jede Reihe

$$\sum_{(l)} l_\beta^2 |a_l|^2 \qquad (\beta = 1, 2, \ldots, n)$$

konvergiert, so konvergiert auch

$$\sum_{(l)} (l_1^2 + \ldots + l_n^2) |a_l|^2 = \sum_{(l)} l^2 |a_l|^2,$$

wobei wir die Summe $l_1^2 + \ldots + l_n^2$ mit l^2 bezeichnet haben.

Wir werden die Konvergenz der Reihe

$$\sum |a_l|^\mu$$

mit Hilfe der bekannten Hölderschen Ungleichung:

$$|\sum u_l v_l| \leqq (\sum |u_l|^p)^{\frac{1}{p}} (\sum |v_l|^q)^{\frac{1}{q}} \qquad \left(p, q > 1, \frac{1}{p} + \frac{1}{q} = 1\right)$$

beweisen. Man setze

$$q = \frac{2}{\mu}, \; p = \frac{1}{1 - \frac{\mu}{2}} = \frac{2}{2 - \mu}, \quad u_l = \frac{1}{l^\mu}, \quad v_l = |l a_l|^\mu.$$

Mit diesen Werten lautet die Höldersche Ungleichung, wie folgt:

$$\sum |a_l|^\mu \leqq \left(\sum \frac{1}{l^{\frac{2\mu}{2-\mu}}}\right)^{\frac{2-\mu}{2}} \left(\sum l^2 |a_l|^2\right)^{\frac{\mu}{2}}.$$

Der zweite Faktor rechts ist konvergent. Der erste ist dann und nur dann konvergent, wenn das Integral

$$\int \frac{1}{l^{\frac{2\mu}{2-\mu}}} dl_1 \ldots dl_n$$

im Unendlichen konvergent ist. Dies gilt dann und nur dann, wenn

$$\frac{2\mu}{2-\mu} > n,$$

also

$$\mu > \frac{2n}{n+2}$$

ist. Die zugehörigen Integrabilitätsexponenten ν sind daher der Ungleichung

$$\nu < \frac{2n}{n-2}$$

unterworfen.

3. Ausdehnung der Gültigkeit unserer Sätze auf den allgemeinen Fall.

Wir setzten beim Beweise unserer Sätze voraus, daß g periodisch für jede ihrer Veränderlichen mit der Periode 2π und T ein Periodenwürfel sei.

Um uns von dieser Voraussetzung zu befreien, führen wir den allgemeinen Fall durch Fortsetzung von g auf den periodischen Fall zurück.

Es ist keine Beschränkung der Allgemeinheit, wenn wir annehmen, daß T im Inneren des Würfels W: $0 \leqq x_\alpha \leqq 2\pi$ liegt, so daß der Abstand zwischen T und dem Rande von W größer, als eine positive Größe d ist. Es sei wieder S die Menge derjenigen Punkte außerhalb T, deren Abstand von T kleiner als d ist. Den Abstand des Punktes $P = (x_1, x_2, \ldots, x_n)$ von T bezeichnen wir mit $t = t(P)$. Wir definieren die Fortsetzungsfunktion, wie folgt:

$$\begin{aligned} G &= g && \text{in } T \\ G &= \left(1 - \frac{t}{d}\right) g && \text{in } S \\ G &= 0 && \text{in } W - S - T. \end{aligned}$$

Wir setzen G außerhalb des Würfels periodisch in Bezug jeder der Veränderlichen mit der Periode 2π fort. Wir wählen $h < d$. Das Integral

$$\int_W (D_\alpha^h G)^2 \, dV$$

bleibt offenbar beschränkt. Es gilt nämlich, wenn wir die charakteristische Funktion des Gebietes $T + S$ mit φ bezeichnen,

$$|D_\alpha^h G(P)| \leqq \varphi(P) \, |D_\alpha^h g(P)| + \left|\frac{g(P)}{d}\right| + \left|\frac{g(P_1)}{d}\right|,$$

wobei wir den Punkt $(x_1, x_2, \ldots, x_\alpha + h, \ldots, x_n)$ durch P_1 bezeichnet haben.

Da die Quadratintegrale der Glieder an der rechten Seite für alle h gleichmäßig beschränkt bleiben, haben wir gezeigt, daß die Voraussetzungen der Sätze aus Abschnitt **1** und **2** für G erfüllt sind. Da g in T mit G identisch ist, so gelten in T für g dieselben Sätze, w. z. b. w.

(Eingegangen am 22. Juni 1939)

Über die Summabilität der Fourierschen Reihe.

Von GÉZA GRÜNWALD in Budapest.

Herrn Prof. Dr. Leopold Fejér im 40. Jahre seiner wissenschaftlichen Tätigkeit.

1. Es sei $f(\xi)$ eine L-integrierbare und nach 2π periodische Funktion der reellen Variabeln ξ und es sei $\{s_n\}$ die Folge der Partialsummen ihrer Fourierschen Reihe an einer Stelle $\xi = x$, wo die Funktion stetig ist. Dann ist nach dem klassischen Fejérschen Satze[1])

$$(1) \qquad \lim_{n=\infty} \frac{s_0+s_1+\ldots+s_n}{n+1} = f(x)$$

und nach HARDY und LITTLEWOOD[2])

$$(2) \qquad \lim_{n=\infty} \frac{s_0^2+s_1^2+\ldots+s_n^2}{n+1} = f^2(x).$$

Vor kurzem hat FEJÉR gefunden[3]), daß auch (2) ein gewöhnlicher Summabilitätssatz ist, nämlich ein Spezialfall des folgenden Summabilitätssatzes, der sich allerdings auf die Fourriersche Reihe einer Funktion zweier Variabeln bezieht:

Es sei $F(\xi, \eta)$ eine L-integrierbare und nach 2π periodische Funktion der reellen Variabeln ξ und η. Bezeichnet dann $\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} A_{mn}$ die Fouriersche Doppelreihe dieser Funktion an einer Stelle $\xi = x$, $\eta = y$, wo die Funktion stetig ist, so ist die einfach unendliche Folge

---

1) L. FEJÉR, Sur les fonctions bornées et intégrables, *Comptes rendus, Paris*, **131** (1900), S. 984—987.

2) G. H. HARDY et J. E. LITTLEWOOD, Sur la série de Fourier d'une fonction à carré sommable, *Comptes rendus, Paris*, **156** (1913), S. 1307—1309.

3) L. FEJÉR, Zur Summabilitätstheorie der Fourierschen und Laplaceschen Reihe, *Proc. Cambridge Phil. Soc.*, **34** (1938), S. 503—509.

$\{s_{nn}\}$ *der Hauptpartialsummen der Reihe* $\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} A_{mn}$ $(C, 3)$*-limitierbar zu* $F(x, y)$ *als Grenzwert; d. h. es ist*

$$(3)\qquad \lim_{n=\infty} \frac{\binom{n+2}{2}s_{00}+\binom{n+1}{2}s_{11}+\ldots+\binom{2}{2}s_{nn}}{\binom{n+3}{3}} = F(x, y),$$

wo

$$s_{nn}=\sum_{\nu=0}^{n}\sum_{\mu=0}^{n} A_{\nu\mu} \qquad (n=0, 1, 2, \ldots).$$

Wir haben neuerdings bewiesen[4]), daß man in dem obigen Satze $(C, 1)$-Summabilität statt $(C, 3)$-Summabilität schreiben kann.

Im Falle $F(\xi, \eta)=f(\xi)f(\eta)$ ergibt sich (2), wegen $s_{nn}=s_n^2$. Unser Satz gilt auch für beliebige höhere Dimensionen, also bleibt (2) auch dann gültig, wenn wir im Exponenten statt 2 ein beliebiges $k>0$ setzen, d. h. es ergibt sich der Carleman—Suttonsche Satz[5]).

Wir betonen, daß *die obige* Fejérsche Summationsmethode, durch die *einfach* unendliche Folge der Hauptpartialsummen, von derjenigen wesentlich *verschieden* ist, die bei der Fourierschen Doppelreihe *gewöhnlich* angewendet wird[6]). Ein großer Vorteil dieser Methode ist neben ihrer Einfachheit, daß sie einen tieferen Einblick in die Theorie der starken Summabilität der Fourierschen Reihe einer Variabeln gewährt, als dies bisher möglich war.

Wir beweisen in dieser Arbeit den folgenden allgemeinen Summabilitätssatz:

Satz A. *Es sei* $F(\xi, \eta)$ *eine L-integrierbare und nach* 2π *periodische Funktion der reellen Variabeln* ξ *und* η. *Bezeichnet dann* $\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} A_{mn}$ *die Fouriersche Doppelreihe dieser Funktion* $F(\xi, \eta)$ *an der Stelle* (ξ, η), *so ist die einfach unendliche Folge* $\{s_{nn}\}$ *der Hauptpartialsummen der Doppelreihe* $\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} A_{mn}$ *fast überall* $(C, 1)$*-limitierbar zu* $F(\xi, \eta)$ *als Grenzwert; d. h. es ist fast überall*

---

[4]) G. Grünwald, Zur Summabilitätstheorie der Fourierschen Doppelreihe, *Proc. Cambridge Phil. Soc.*, 35 (1939), S. 343—350.

[5]) Siehe z. B. A. Zygmund, *Trigonometrical Series* (Warsaw, 1935), S. 238.

[6]) Diesbezüglich siehe L. Fejér [3]), insb. S. 504, Fußnote.

im Quadrat $-\pi \leqq \xi \leqq \pi, \quad -\pi \leqq \eta \leqq \pi$

$$\lim_{n=\infty} \frac{s_{00}+s_{11}+\ldots+s_{nn}}{n+1} = F(\xi, \eta), \tag{4}$$

wo

$$s_{nn} = \sum_{\nu=0}^{n} \sum_{\mu=0}^{n} A_{\nu\mu} \qquad (n=0, 1, 2, \ldots).$$

Auch dieser Satz bleibt im Falle mehrerer Variabeln gültig und der Beweis verläuft analog wie für Satz A; deshalb verzichten wir hier auf eine nähere Ausführung.

Als Korollar des Satzes A ergibt sich der folgende allgemeine starke Summabilitätssatz für gewöhnliche Fouriersche Reihen:

Satz B. *Es seien $f(\xi)$ und $g(\eta)$ L-integrierbare und nach 2π periodische Funktionen der reellen Variabeln ξ bzw. η. Dann ist für fast alle (ξ, η) im Quadrat* $-\pi \leqq \xi \leqq \pi, \quad -\pi \leqq \eta \leqq \pi$

$$\lim_{n=\infty} \frac{s_0[f(\xi)]s_0[g(\eta)]+s_1[f(\xi)]s_1[g(\eta)]+\ldots+s_n[f(\xi)]s_n[g(\eta)]}{n+1} = \tag{5}$$
$$= f(\xi)g(\eta),$$

wo $s_\nu[f(\xi)]$, $s_\nu[g(\eta)]$ die Partialsummen der Fourierschen Reihen der Funktionen $f(\xi)$ bzw. $g(\eta)$ an den Stellen ξ bzw. η bezeichnen.

Es ist naheliegend zu fragen, ob man im Spezialfall $F(\xi, \eta) = f(\xi)f(\eta)$ bzw. $F(\xi_1, \xi_2, \ldots, \xi_p) = f(\xi_1)f(\xi_2)\ldots f(\xi_p)$ die mehrdimensionale Ausnahmenullmenge des Satzes A auf eine eindimensionale Nullmenge reduzieren kann und so der Hardy—Littlewoodsche bzw. Carleman—Suttonsche Satz aus der einzigen Bedingung folgt, daß $f(\xi)$ L-integrierbar ist[7]). Wir hoffen auf diesen Gegenstand zurückkehren zu können.

2. Die folgende Verschärfung des Vitalischen Überdeckungsatzes stammt von J. Marcinkiewicz und A. Zygmund[8]):

Lemma 1. *Es sei $\alpha > 0$ eine feste Zahl und E eine Punktmenge in der Ebene, für welche $0 < mE < \infty$ ist. Es sei jedem*

---

[7]) Den Hardy—Littlewoodschen Satz hat unter dieser allgemeinen Voraussetzung J. Marcinkiewicz bewiesen: Sur la sommabilité forte de séries de Fourier, *Journal London Math. Soc.*, 14 (1939), S. 162—168. Auch bei ihm bildet den Ausgangpunkt die wichtige Formel von Fejér (S. 163, Formel 2. 4 bei Marcinkiewicz), welche die fraglichen Mittel in Form eines *Doppelintegrales mit nichtnegativem Kern* darstellt.

[8]) J. Marcinkiewicz and A. Zygmund, On the Summability of double Fourier series, *Fundamenta Math.*, 32 (1939), S. 122—132, insb. Lemma 1.

Punkt P von E ein Rechteck R(P) zugeordnet, dessen Seiten von der Länge $\delta(P)$ bzw. $\alpha\delta(P)$ achsenparallel sind und dessen Zentrum in P liegt. Dann gibt es endlichviele einander nicht überdeckende Rechtecke so, daß[9])

$$\sum_{\nu=0}^{n} mR(P_\nu) > cmE.$$

Aus diesem Lemma folgt[10])

Lemma 2. *Es sei $F(\xi, \eta)$ im Quadrat $Q': -2\pi \leqq \xi,\ \eta \leqq 2\pi$ integrierbar und es sei $\alpha > 0$ eine feste Zahl. Man setze für einen beliebigen Punkt (ξ, η) im Quadrat $Q: -\pi \leqq \xi,\ \eta \leqq \pi$*

$$F_\alpha^*(\xi, \eta) = \operatorname{Max}_h \frac{1}{4\alpha h^2} \int_{-\alpha h}^{\alpha h} dt \int_{-h}^{h} |F(\xi + u, \eta + t)|\, du, \tag{6}$$

wo h so klein ist, daß das Integrationsgebiet in Q' liege. Ferner sei

$$F^*(\xi, \eta) = \operatorname{Max}_s \{F_{2^s}^*(\xi, \eta) 2^{-|s|/2}\} \qquad (s = 0, \pm 1, \pm 2, \ldots). \tag{7}$$

Dann ist

$$mE\{F^*(\xi, \eta) > \omega\} < \frac{c}{\omega} \iint_{(Q')} |F(\xi, \eta)|\, d\xi\, d\eta, \tag{8}$$

wo, wie üblich, $E\{F^*(\xi, \eta) > \omega\}$ die Menge der Punkte (ξ, η) bezeichnet, für welche $F^*(\xi, \eta) > \omega$.

Auf dieselbe Art wie Lemma 1 bzw. der Vitalische Überdeckungsatz, läßt sich beweisen

Lemma 3. *Es sei $\alpha > 0$ eine feste Zahl und E eine Punktmenge in der Ebene, für welche $0 < mE < \infty$ ist. Jedem Punkte P von E sei ein Parallelogramm R(P) mit dem Zentrum in P und mit den Seitenlängen $\delta(P)$ bzw. $\alpha\delta(P)$ so zugeordnet, daß die eine Seite von R parallel zur y-Achse und der Winkel zwischen der anderen Seite und der x-Achse 45° ist. Dann gibt es endlichviele einander nicht überdeckende Parallelogramme so, daß*

$$\sum_{\nu=0}^{n} mR(P_\nu) > cmE.$$

Aus diesem Lemma folgt (so, wie Lemma 2 aus Lemma 1, vgl.[10]))

[9]) c ist eine absolute Konstante (unabhängig auch von α). Im Folgenden bezeichnen wir einfachheitshalber die verschiedenen absoluten Konstanten mit c, also ohne Indizes.

[10]) Siehe [8]), Lemma 2 und Lemma 3.

Lemma 4. *Es sei $F(\xi, \eta)$ im Quadrat $Q'' : -3\pi \leqq \xi, \eta \leqq 3\pi$ L-integrierbar und es sei $\alpha > 0$ eine feste Zahl. Man setze für einen beliebigen Punkt (ξ, η) im Quadrat $Q : -\pi \leqq \xi, \eta \leqq \pi$*

$$F_{\alpha}^{**}(\xi, \eta) = \operatorname*{Max}_{h} \frac{1}{4\alpha h^2} \int_{-\alpha h}^{\alpha h} dt \int_{-h+t}^{h+t} |F(\xi + u, \eta + t)| du, \tag{9}$$

wo h so klein ist, daß das Integrationsgebiet in Q'' liege. Ferner sei

$$F^{**}(\xi, \eta) = \operatorname*{Max}_{h} \{F_{2^s}^{**}(\xi, \eta) 2^{-|s|/2}\} \qquad (s = 0, \pm 1, \pm 2, \ldots). \tag{10}$$

Dann ist

$$mE\{F^{*}(\xi, \eta) > \omega\} < \frac{c}{\omega} \iint_{(Q'')} |F(\xi, \eta) d\xi d\eta. \tag{11}$$

3. Für den Beweis unseres Satzes A ist es hinreichend zu zeigen, daß

$$\sigma^*(\xi, \eta; F) = \operatorname*{Max}_{n} |\sigma_n(\xi, \eta; F)| \leqq c(F^*(\xi, \eta) + F^{**}(\xi, \eta)),$$

wo

$$\sigma_n(\xi, \eta; F) = \frac{s_{00} + s_{11} + \ldots + s_{n-1, n-1}}{n}.$$

Ist nämlich $F(\xi, \eta)$ eine periodische Funktion, dann folgt aus (8) und (11), daß

$$mE\{\sigma^*(\xi, \eta; F) > \omega\} > \frac{c}{\omega} \iint_{(Q)} |F(\xi, \eta)| d\xi d\eta. \tag{12}$$

Sei $\varepsilon > 0$ beliebig und $F = F_1 + F_2$, wo F_1 ein trigonometrisches Polynom ist und F_2 eine Funktion, für welche

$$mE\{\sigma^*(\xi, \eta; F_2) > \varepsilon\} < \varepsilon \tag{13}$$

(eine solche Zerlegung von F ist wegen (12) möglich). Da $\sigma_n(\xi, \eta; F_1)$ gleichmäßig gegen F_1 konvergiert und da $|\sigma_n(\xi, \eta; F_2)| < \varepsilon$ außer höchstens auf eine Menge vom Maße $< \varepsilon$, so folgt, daß $\sigma_n(\xi, \eta; F)$ fast überall gegen F konvergiert.

4. Die explizite Form von $\sigma_n(\xi, \eta; F)$ ist die Folgende:

$$\sigma_n(\xi, \eta; F) =$$

$$= \frac{1}{4\pi^2 n} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} F(u + \xi, t + \eta) \sum_{k=0}^{n} \frac{\sin(k - 1/2) u \sin(k - 1/2) t}{\sin u/2 \sin t\ 2} du\, dt =$$

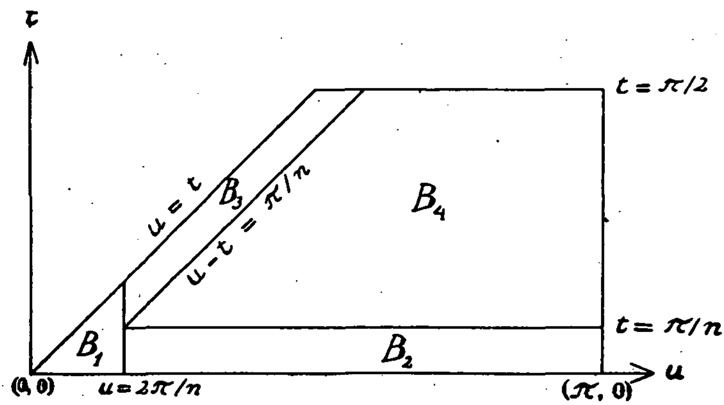
(14)

$$= \frac{1}{\pi^2} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} F(u + \xi, t + \eta) k_n(u, t) du\, dt.$$

Eine einfache Rechnung ergibt, daß

$$(15)\quad k_n(u,t)=\frac{1}{16n\sin{}^u\!/_2\sin{}^t\!/_2}\left(\frac{\sin n(u-t)}{\sin{}^1\!/_2(u-t)}-\frac{\sin n(u+t)}{\sin{}^1\!/_2(u+t)}\right).$$

Wir werden $k_n(u, t)$ in verschiedenen, in Q liegenden Bereichen abschätzen. Aus Symmetriegründen beschränken wir uns auf den Fall $0\leqq t\leqq u\leqq \pi$, $t\leqq {}^\pi\!/_2$. Wir zerlegen diesen Bereich B, wie es aus der Figur ersichtlich ist, in vier Teile.



a) In B_1 folgt aus dem Ausdruck von $k_n(u, t)$ in (14)

$$(16)\qquad |k_n(u,t)|\leqq\frac{1}{16n}\sum_{k=1}^{n}(2k-1)^2=cn^2.$$

b) In B_2 ist wegen $u-t\geqq{}^\pi\!/_n$ und $u-t\geqq{}^u\!/_2$

$$\frac{\sin n(u-t)}{\sin{}^1\!/_2(u-t)}-\frac{\sin n(u+t)}{\sin{}^1\!/_2(u+t)}=-2\frac{d}{dx}\left(\frac{\sin nx}{\sin{}^x\!/_2}\right)_{x^*},\ u-t<x^*<u+t,$$

also

$$\left|\frac{\sin n(u-t)}{\sin{}^1\!/_2(u-t)}-\frac{\sin n(u+t)}{\sin{}^1\!/_2(u+t)}\right|<ct\left(\frac{n}{u-t}+\frac{1}{(u-t)^2}\right)<c\left(\frac{2nt}{u}+\frac{4}{u^2}\right)$$

und nach (15)

$$(17)\qquad |k_n(u,t)|\leqq c\frac{1}{nut}\,n\frac{t}{u}=c\frac{1}{u^2}.$$

c) In B_3 ist

$$(18)\qquad |k_n(u,t)|\leqq c\frac{1}{n}\frac{1}{ut}n=c\frac{1}{ut}.$$

d) In B_4 ist

$$|k_n(u, t)| \leq c \frac{1}{n} \frac{1}{ut(u-t)}. \tag{19}$$

5. Es sei

$$G_n(\xi, \eta) = \frac{1}{\pi^2} \iint\limits_{(B)} F(\xi+u, \eta+t)\, k_n(u, t)\, du\, dt \tag{20}$$

und p eine natürliche Zahl mit $2^p \leq n < 2^{p+1}$. Dann ist nach den Abschätzungen des vorigen Paragraphen

$$|G_n(\xi, \eta)| \leq c\Bigg(2^{2p} \iint\limits_{(B_1)} |F(\xi+u, \eta+t)|\, du\, dt + \iint\limits_{(B_2)} \frac{|F(\xi+u, \eta+t)|}{u^2} du\, dt +$$

$$+ \iint\limits_{(B_3)} \frac{|F(\xi+u, \eta+t)|}{ut} du\, dt + 2^{-p} \iint\limits_{(B_4)} \frac{|F(\xi+u, \eta+t)|}{ut(u-t)} du\, dt\Bigg) \leq$$

$$\leq c\Bigg(2^{2p} \int_0^{\pi 2^{-p}} \int_0^{\pi 2^{-p}} |F(\xi+u, \eta+t)|\, du\, dt + \int_{\pi 2^{-p}}^{\pi} \int_0^{\pi 2^{-p}} \frac{|F(\xi+u, \eta+t)|}{u^2} du\, dt + \tag{21}$$

$$+ \int_{\pi 2^{-p}}^{\pi} du \int_{u-\pi 2^{-p}}^{u} \frac{|F(\xi+u, \eta+t)|}{ut} du\, dt + 2^{-p} \int_{\pi 2^{-p}}^{\pi} du \int_{\pi 2^{-p-1}}^{u-\pi 2^{-p-1}} \frac{|F(\xi+u, \eta+t)|}{ut(u-t)} du\, dt\Bigg) =$$

$$= c\,(A_p(\xi, \eta) + B_p(\xi, \eta) + C_p(\xi, \eta) + D_p(\xi, \eta)).$$

Es sei

$$A^*(\xi, \eta) = \operatorname*{Max}_p A_p(\xi, \eta) \qquad p = 1, 2, \ldots \tag{22}$$

und $B^*(\xi, \eta)$, $C^*(\xi, \eta)$, $D^*(\xi, \eta)$ haben eine analoge Deutung. Wir bewiesen nun, daß

$$mE\{A^*(\xi, \eta) > \omega\} < \frac{c}{\omega} \iint\limits_{(Q)} |F(\xi, \eta)|\, d\xi\, d\eta \tag{23}$$

und die analogen Abschätzungen für $B^*(\xi, \eta)$, $C^*(\xi, \eta)$, $D^*(\xi, \eta)$. Nach (21) ergibt sich, wenn wir die Bezeichnungen von Lemma 2 anwenden,

$$A_p(\xi, \eta) \leq c\, 2^{2p}\, 2^{-2p} F_1^*(\xi, \eta) < c F^*(\xi, \eta), \tag{24}$$

$$
(25)\quad \begin{aligned}
B_p(\xi,\eta) &= \sum_{i=0}^{p-1} \int_0^{\pi 2^{-p}} dt \int_{\pi 2^{-i-1}}^{\pi 2^{-i}} \frac{|F(\xi+u,\eta+t)|}{u^2}\, du \leqq \\
&\leqq c \sum_{i=0}^{p-1} 2^{2i} \int_{-\pi 2^{-p}}^{\pi 2^{-p}} dt \int_{-\pi 2^{-i}}^{\pi 2^{-i}} |F(\xi+u,\eta+t)|\, du \leqq \\
&\leqq \sum_{i=1}^{p-1} 2^{2i} 2^{-(p+i)} F^*_{2^{i-p}}(\xi,\eta) \leqq c \sum_{i=0}^{p-1} 2^{i-p+|i-p|/2} F^*(\xi,\eta) = \\
&= cF^*(\xi,\eta) \sum_{i=0}^{p-1} 2^{-(p-i)/2} = cF^*(\xi,\eta).
\end{aligned}
$$

Ferner ergibt sich nach (21), wenn wir die Bezeichnungen von Lemma 4 anwenden

$$
(26)\quad \begin{aligned}
C_p(\xi,\eta) &= \sum_{i=1}^{p-2} \int_{\pi 2^{-i-1}}^{\pi 2^{-i}} du \int_{u-\pi 2^{-p}}^{u} \frac{|F(\xi+u,\eta+t)|}{ut}\, dt + \\
&\qquad + \int_{\pi 2^{-p}}^{\pi 2^{-p+1}} du \int_{u-\pi 2^{-p}}^{u} \frac{|F(\xi+u,\eta+t)|}{ut}\, dt \leqq \\
&\leqq c \sum_{i=0}^{p-2} \frac{2^i}{2^{-i-1}-2^{-p}} \int_{-\pi 2^{-i}}^{\pi 2^{-i}} du \int_{u-\pi 2^{-p}}^{u+\pi 2^{-p}} |F(\xi+u,\eta+t)|\, dt + \\
&\qquad + c\, 2^{2p} \int_{-\pi 2^{-p}}^{\pi 2^{-p}} du \int_{u-\pi 2^{-p}}^{u+\pi 2^{-p}} |F(\xi+u,\eta+t)|\, dt \leqq \\
&\leqq c \sum_{i=0}^{p-1} \frac{2^i 2^{-(i+p)}}{2^{-i-1}-2^{-p}} F^{**}_{2^{i-p}}(\xi,\eta) + c\, 2^{2p} 2^{-2p} F_1^{**}(\xi,\eta) \leqq \\
&\leqq c \sum_{i=0}^{p-1} \frac{2^{-p} 2^{|i-p|/2}}{2^{-i-1}-2^{-p}} F^{**}(\xi,\eta) + cF^{**}(\xi,\eta) = cF^{**}(\xi,\eta),
\end{aligned}
$$

$$
\begin{aligned}
D_p(\xi,\eta) &= 2^{-p} \sum_{i=0}^{p-1} \int_{\pi 2^{-i-1}}^{\pi 2^{-i}} du \left(\sum_{j=0}^{p-i-2} \int_{u-\pi 2^{j-p}}^{u-\pi 2^{j-p-1}} \frac{|F(\xi+u,\eta+t)|}{ut(u-t)}\, dt \right) + \\
&\qquad + 2^{-p} \sum_{i=0}^{p-1} \int_{\pi 2^{-i-1}}^{\pi 2^{-i}} du \int_{\pi 2^{-p-1}}^{u-\pi 2^{-i-2}} \frac{|F(\xi+u,\eta+t)|}{ut(u-t)}\, dt = I_1 + I_2;
\end{aligned}
$$

$$I_1 \leqq c\,2^{-p} \sum_{i=0}^{p-1} \int_{\pi 2^{-i-1}}^{\pi 2^{-i}} du \left(\sum_{j=0}^{p-i-2} \frac{2^{p+i-j+1}}{2^{-i-1}-2^{j-p}} \int_{u-\pi 2^{j-p}}^{u-\pi 2^{j-p-1}} |F(\xi+u, \eta+t)|\, dt \right) \leqq$$

$$\leqq c \sum_{i=0}^{p-1} \left(\sum_{j=0}^{p-i-2} \frac{2^{-p}}{2^{-i-1}-2^{j-p}} F^{**}_{2^{j-p-i}}(\xi, \eta) \right) \leqq$$

$$\leqq c F^{**}(\xi, \eta) \sum_{i=0}^{p-1} \left(\sum_{j=0}^{p-i-2} \frac{2^{-p+|j-p+i|/2}}{2^{-i-1}-2^{j-p}} \right) \leqq$$

$$\leqq c F^{**}(\xi, \eta) \sum_{i=0}^{p-1} \left(\sum_{j=0}^{p-i-2} 2^{-(p-i)/2-j/2} \right) = c F^{**}(\xi, \eta),$$

$$I_2 \leqq c\,2^{-p} \sum_{i=0}^{p-1} \int_{\pi 2^{-i-1}}^{\pi 2^{-i}} du \left(\sum_{j=0}^{p-i} 2^{i} \int_{\pi 2^{-i-j+1}}^{\pi 2^{-i-j}} \frac{|F(\xi+u, \eta+t)|}{ut}\, dt \right) \leqq$$

$$\leqq c \sum_{i=0}^{p-1} \left(\sum_{j=0}^{p-i} 2^{i-p} F^{**}_{2^j}(\xi, \eta) \right) \leqq c F^{**}(\xi, \eta) \sum_{i=0}^{p-1} \left(\sum_{j=0}^{p-i} 2^{i-p+j/2} \right) \leqq$$

$$\leqq c F^{**}(\xi, \eta) \sum_{i=0}^{p-1} 2^{-(p-i)/2} = c F^{**}(\xi, \eta),$$

$$D_p(\xi, \eta) = I_1 + I_2 \leqq c F^{**}(\xi, \eta). \tag{27}$$

Die Ungleichungen (24), (25) mit Lemma 2 und die Ungleichungen (26), (27) mit Lemma 4 ergeben die Ungleichungen (23). Daraus folgt, daß

$$|G_n(\xi, \eta)| \leqq c(F^{**}(\xi, \eta) + F^{**}(\xi, \eta)).$$

Da $\sigma_n(\xi, \eta; F)$ die Summe von mit (21) analogen Integralen ist, ergibt sich die Ungleichung

$$|\sigma_n(\xi, \eta; F)| \leqq c(F^{*}(\xi, \eta) + F^{**}(\xi, \eta)),$$

womit unser Satz bewiesen ist.

(Eingegangen am 10. Januar 1940.)

Über Integralungleichungen zwischen einer Funktion und ihrer Ableitung.

Von BÉLA V. SZ. NAGY in Szeged.

§ 1. Die Ungleichungen.

Es sei $y(x)$ eine auf $(-\infty, \infty)$ definierte totalstetige Funktion, die also die Integralfunktion ihrer (fast überall existierenden, nach Lebesgue integrierbaren) Derivierten $y'(x)$ ist.

Sind die Integrale

$$J_\alpha = \int_{-\infty}^{\infty} |y|^\alpha dx \quad \textit{und} \quad K_p = \int_{-\infty}^{\infty} |y'|^p dx$$

für ein bestimmtes $\alpha > 0$ und für ein bestimmtes $p \geqq 1$ beide endlich, so hat man

$$(1) \qquad \max_{-\infty < x < \infty} |y| \leqq \left(\frac{q}{2}\right)^{\frac{1}{q}} J_\alpha^{\frac{p-1}{pq}} K_p^{\frac{1}{pq}} \qquad \left(q = 1 + \frac{p-1}{p}\alpha\right).$$

Für $\beta > 0$ hat man ferner

$$(2) \qquad J_{\alpha+\beta} \leqq \left[\frac{q}{2} H\left(\frac{q}{\beta}, \frac{p-1}{p}\right)\right]^{\frac{\beta}{q}} J_\alpha^{1+\beta\frac{p-1}{pq}} K_p^{\frac{\beta}{pq}},$$

wo

$$H(u, v) = \frac{(u+v)^{-(u+v)}\Gamma(1+u+v)}{u^{-u}\Gamma(1+u)v^{-v}\Gamma(1+v)}, \quad H(u,0) = H(0,v) = 1.\ ^{1)}$$

---

1) Dies ist für $u \geqq 0$, $v \geqq 0$ eine monoton fallende Funktion ihrer beiden Veränderlichen. Da nun $H(u,0) = 1$ und $H(u,1) = \left(1 + \frac{1}{u}\right)^{-u}$, so hat man $1 > H(u,v) > \left(1+\frac{1}{u}\right)^{-u} > \frac{1}{e}$ für $0 < v < 1$. Es gilt also

$$1 \geqq H\left(\frac{q}{\beta}, \frac{p-1}{p}\right) > \left(1 + \frac{\beta}{q}\right)^{-\frac{q}{\beta}} > \frac{1}{e}.$$

Die Funktion $H(u,v)$ wurde von Herrn E. SCHMIDT eingeführt und

Die angeführten Konstanten sind genau.

Schreibt man (2) in der Form

$$(3) \qquad \left(\frac{J_{\alpha+\beta}}{J_\alpha}\right)^{\frac{1}{\beta}} \leqq \left[\frac{q}{2} H\left(\frac{q}{\beta}, \frac{p-1}{p}\right)\right]^{\frac{1}{q}} J_\alpha^{\frac{p-1}{pq}} K_p^{\frac{1}{pq}}$$

und läßt man β über alle Grenzen wachsen, so erhält man im Grenzfall die Ungleichung (1).

Im Falle $p=1$ gilt in (1) das Gleichheitszeichen immer dann, wenn $|y(x)|$ bis zu einem Wert x_0 von x monoton wächst, dann aber monoton fällt. Im Falle $p>1$ gilt in (1) das Gleichheitszeichen dann und nur dann, wenn $y(x)$ die Form $a y_{p\alpha}(|bx+c|)$ hat, wo a, b, c beliebige Konstanten sind $(b \neq 0)$ und $y_{p\alpha}(t)$ folgendermaßen definiert ist:

$$\text{im Falle } \alpha < p: \quad y_{p\alpha}(t) = \begin{cases} (1-t)^{\frac{p}{p-\alpha}} & \text{für } 0 \leqq t \leqq 1, \\ 0 & \text{für } t > 1, \end{cases}$$

$$\text{im Falle } \alpha = p: \quad y_{p\alpha}(t) = e^{-t} \qquad \text{für } t \geqq 0,$$

$$\text{im Falle } \alpha > p: \quad y_{p\alpha}(t) = (1+t)^{\frac{p}{p-\alpha}} \qquad \text{für } t \geqq 0.$$

In (2) ist im Falle $p=1$ immer (wenn $y \not\equiv 0$) das Zeichen „$<$" gültig. Im Falle $p>1$ gilt in (2) das Gleichheitszeichen dann und nur dann, wenn $y(x)$ die Form $a y_{p\alpha\beta}(|bx+c|)$ hat, wo a, b, c beliebige Konstanten sind $(b \neq 0)$ und $y_{p\alpha\beta}(t)$ für $t \geqq 0$ folgendermaßen definiert ist: $u = y_{p\alpha\beta}(t)$ ist die inverse Funktion der Funktion

$$(4) \qquad t = \int_u^1 \frac{ds}{[s^\alpha(1-s^\beta)]^{\frac{1}{p}}} \qquad (0 \leqq u \leqq 1);$$

$y_{p\alpha\beta}(t)$ ist also eine monoton fallende Funktion, sie ist im Falle

---

untersucht in seiner Arbeit: Über die Ungleichung, welche die Integrale über eine Potenz einer Funktion und über eine andere Potenz ihrer Ableitung verbindet, *Math. Annalen*, **117** (1940), S. 301–326, wo er Ungleichungen vom Typus $\left\{\int_0^1 |y|^a\,dx\right\}^{\frac{1}{a}} \leqq C \left\{\int_0^1 |y'|^b\,dx\right\}^{\frac{1}{b}}$ $(a>0,\ b\geqq 1)$ z. B. unter der Bedingung: $\max y + \min y = 0$ aufstellt. Wegen ihrer Inhomogenität gegenüber der Substitution $x \to \varrho x$ kann eine solche Ungleichung im Falle eines unendlichen Integrationsintervalls nicht mehr bestehen. Im Beweis von (2) werden wir uns teils der Methode des Herrn Schmidt bedienen.

$\alpha \geqq p$ immer positiv, im Falle $\alpha < p$ wird sie dagegen für

$$t_0 = \int_0^1 \frac{ds}{[s^\alpha(1-s^\beta)]^{\frac{1}{p}}}$$

gleich 0; wir wollen in diesem Falle $y_{p\alpha\beta}(t) = 0$ auch für $t > t_0$ setzen.

Im Falle $\alpha = p = 2$ ergibt (1) insbesondere:

$$y(0) \leqq J_2^{\frac{1}{4}} K_2^{\frac{1}{4}}.$$

Wendet man diese Ungleichung auf die Cosinus-Transformierte $\sqrt{\frac{2}{\pi}} \int_0^\infty \cos tx \,.\, g(t)\, dt$ einer quadratisch integrierbaren reellwertigen Funktion $g(t)$ an, für die auch $t\, g(t)$ quadratisch integrierbar ist, so erhält man auf Grund des Parsevalschen Satzes:

$$\int_0^\infty g(t)\, dt \leqq \sqrt{\pi} \left\{ \int_0^\infty g^2(t)\, dt \int_0^\infty t^2 g^2(t)\, dt \right\}^{\frac{1}{4}}$$

Dies ist eine bekannte Ungleichung von CARLSON[2]). Gleichheit tritt hier nur für die Cosinus-Transformierten der Funktionen $a\, e^{-|bx|}$ auf.

Schreiben wir nun (2) für $\alpha = \beta = 1$ und $p = 2$ auf. Dann wird $q = \frac{3}{2}$ und eine leichte Rechnung zeigt, daß $H\left(\frac{3}{2}, \frac{1}{2}\right) = \frac{\sqrt{3}}{\pi}$. Also hat man

$$J_2 \leqq \frac{3}{(4\pi)^{\frac{2}{3}}} J_1^{\frac{4}{3}} K_2^{\frac{1}{3}}. \tag{5}$$

Da, wie man leicht nachrechnet,

$$y_{211}(t) = \begin{cases} \cos^2 \frac{t}{2} & \text{für } 0 \leqq t \leqq \pi, \\ 0 & \text{für } t > \pi, \end{cases}$$

so wird in (5) Gleichheit dann und nur dann bestehen, wenn

---

[2]) Siehe F. CARLSON, Sur une inégalité, *Arkiv för Matematik, Astronomi och Fysik*, 25 B (1934), No. 1.

$$y = \begin{cases} a\cos^2(bx+c) & \text{für} \quad |bx+c| \leqq \frac{\pi}{2} \\ 0 & \text{für} \quad |bx+c| > \frac{\pi}{2} \end{cases}$$

(a, b, c sind beliebige Konstanten, $b \neq 0$).

Es sei nun $f(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{itx} y(x)\, dx$; man hat wegen des Parsevalschen Satzes:

$$J_2 = \int\limits_{-\infty}^{\infty} |f(t)|^2\, dt, \quad K_2 = \int\limits_{-\infty}^{\infty} t^2 |f(t)|^2\, dt.$$

Ist $y(x)$ überall nichtnegativ, so ist $J_1 = \int\limits_{-\infty}^{\infty} y(x)\, dx = \sqrt{2\pi}\, f(0)$. (5) ergibt also:

$$\int\limits_{-\infty}^{\infty} |f(t)|^2\, dt \leqq \frac{3}{(4\pi)^{\frac{2}{3}}} \left(\sqrt{2\pi} f(0)\right)^{\frac{4}{3}} \left(\int\limits_{-\infty}^{\infty} t^2 |f(t)|^2\, dt \right)^{\frac{1}{3}},$$

oder

$$f(0) \geqq \left(\frac{4}{27}\right)^{\frac{1}{4}} \frac{\left\{ \int\limits_{-\infty}^{\infty} |f(t)|^2\, dt \right\}^{\frac{3}{4}}}{\left\{ \int\limits_{-\infty}^{\infty} t^2 |f(t)|^2\, dt \right\}^{\frac{1}{4}}},$$

gültig für alle komplexwertigen Funktionen, für die die rechtsstehenden Integrale endlich sind und *deren Fouriersche Transformierte überall nichtnegativ reellwertig sind.*

Die Ungleichungen (1), (2) lassen sich auch auf den Fall übertragen, daß die Funktionen auf einem endlichen Intervall definiert sind, vorausgesetzt, daß sie in mindestens einem Punkte des Intervalles verschwinden. Auf diese Frage wollen wir aber nicht näher eingehen.

§ 2. Beweis der ersten Ungleichung.

Es folgt aus der Endlichkeit von J_α, daß es gegen ∞ bzw. $-\infty$ strebende Folgen a_n, b_n gibt derart, daß $y(a_n) \to 0$, $y(b_n) \to 0$.

Dann ist aber im Falle $p = 1$:

$$\int\limits_{-\infty}^{\infty} |y'|\, dx \geqq \mp \int\limits_{\xi}^{\infty} y'\, dx \pm \int\limits_{-\infty}^{\xi} y'\, dx = \lim_{n \to \infty} \left(\mp \int\limits_{\xi}^{a_n} \pm \int\limits_{b_n}^{\xi} \right) y'\, dx = \pm 2y(\xi),$$

also $|y(\xi)| \leqq \frac{1}{2} K_1$, wie auch ξ gewählt wurde. Damit ist aber (1) im Falle $p = 1$ bewiesen; man sieht auch, daß Gleichheitszeichen darin dann und nur dann gilt, wenn $|y(x)|$ bis zu einer gewissen Stelle monoton wächst, dann monoton fällt.

Es sei nun $p > 1$. Anwendung der Hölderschen Ungleichung ergibt:

(6) $$\int_{-\infty}^{\infty} |y|^{\frac{p-1}{p}\alpha} |y'|\, dx \leqq \left(\int_{-\infty}^{\infty} |y|^{\alpha}\, dx\right)^{\frac{p-1}{p}} \left(\int_{-\infty}^{\infty} |y'|^{p}\, dx\right)^{\frac{1}{p}} = J_\alpha^{\frac{p-1}{p}} K_p^{\frac{1}{p}}.$$

Weiter hat man

$$\int_{-\infty}^{\infty} |y|^{\frac{p-1}{p}\alpha} |y'|\, dx \geqq \left[-\int_{0}^{\infty} + \int_{-\infty}^{0}\right] (\operatorname{sgn} y)\, |y|^{\frac{p-1}{p}\alpha}\, y'\, dx =$$

(7) $$= \frac{1}{1+\frac{p-1}{p}\alpha}\left[-\int_{0}^{\infty} + \int_{-\infty}^{0}\right] \frac{d}{dx} |y|^{1+\frac{p-1}{p}\alpha}\, dx =$$

$$= \frac{1}{q}\left[2|y(0)|^q - \lim |y(a_n)|^q - \lim |y(b_n)|^q\right] = \frac{2}{q} |y(0)|^q.$$

Aus (6) und (7) folgt:

(8) $$|y(0)|^q \leqq \frac{q}{2} J_\alpha^{\frac{p-1}{p}} K_p^{\frac{1}{p}}.$$

Wenden wir (8) auf die Funktion $y_\xi(x) = y(x+\xi)$ (ξ beliebig) an, so ergibt sich:

$$|y(\xi)|^q \leqq \frac{q}{2} J_\alpha^{\frac{p-1}{p}} K_p^{\frac{1}{p}},$$

womit (1) bewiesen ist.

Wenden wir jetzt (1) auf die Funktion $\bar{y}_\lambda(x) = y(|x|+\lambda)$ ($\lambda > 0$) statt $y(x)$ an, so ergibt sich:

$$\max_{-\infty < x < \infty} |\bar{y}_\lambda|^q = \max_{x \geqq \lambda} |y|^q \leqq \frac{q}{2}\left(2\int_{\lambda}^{\infty} |y|^{\alpha} dx\right)^{\frac{p-1}{p}} \left(2\int_{\lambda}^{\infty} |y'|^{p} dx\right)^{\frac{1}{p}},$$

woraus folgt, daß $\lim\limits_{x\to\infty} y(x) = 0$. Ebenso beweist man, daß $\lim\limits_{x\to-\infty} y(x) = 0$. Die Funktion $|y(x)|$ nimmt daher ihr Maximum an.

Da (1) gegenüber einer Translation $x \to x+\xi$ invariant ist und da das Ersetzen von y durch ϱy das Multiplizieren beider Seiten durch $|\varrho|$ bewirkt, darf man sich beim Aufsuchen der nichttrivialen Extremalen auf solche Funktionen $y(x)$ beschränken, für welche $y(0)=\max|y|=1$. Gleichheit in (1) wird dann Gleichheit in (8) bedeuten; in (8) besteht aber das Gleichheitszeichen dann und nur dann, wenn es auch in (6) und (7) gleichzeitig besteht. In (6) gilt aber Gleichheit dann und nur dann, wenn $|y|^\alpha$ und $|y'|^p$ proportional sind, d. h., wenn fast überall

$$(9) \qquad |y'(x)|=\varrho|y(x)|^{\frac{\alpha}{p}}$$

mit einer positiven Konstanten ϱ gilt. In (7) aber wird das Gleichheitszeichen dann und nur dann bestehen, wenn fast überall, wo $y(x)\neq 0$:

$$(10) \qquad \begin{aligned} |y'(x)|&=\operatorname{sgn} y(x)\,.\,y'(x) \quad \text{für} \quad x<0, \\ -|y'(x)|&=\operatorname{sgn} y(x)\,.\,y'(x) \quad \text{für} \quad x>0. \end{aligned}$$

Letzte Bedingung ist nun offenbar damit gleichbedeutend, daß $[y^2(x)]'=2y(x)y'(x)$ für $x<0$ nichtnegativ, für $x>0$ nichtpositiv ausfällt, also, daß $y^2(x)$ für $x<0$ monoton wächst, für $x>0$ monoton fällt. Dies ist aber wegen $y(0)>0$ dann und nur dann der Fall, wenn $y(x)$ ebenfalls für $x<0$ monoton wächst, für $x>0$ monoton fällt und $\geqq 0$ ist. Für diese Funktionen nimmt (9) die folgende Gestalt an:

$$y'(x)=\varrho[y(x)]^{\frac{\alpha}{p}} \text{ für } x<0, \quad -y'(x)=\varrho[y(x)]^{\frac{\alpha}{p}} \text{ für } x>0.$$

Diese Differentialgleichung hat aber unter der Bedingung $y(0)=1$ die einzige Lösung $y=y_{p\alpha}(|\sigma x|)$, wo $y_{p\alpha}(t)$ die in § 1 definierte Funktion ist und $\sigma=\varrho\left(1-\frac{\alpha}{p}\right)$ bzw. $\sigma=\varrho$, je nachdem $\alpha\neq p$ oder $\alpha=p$, gesetzt wird. Die allgemeine Extremale erhält man aus dieser Funktion durch Multiplizieren mit Konstanten und Verschiebungen längs der x-Achse.

§ 3. Beweis der zweiten Ungleichung.

Im Falle $p=1$ haben wir wegen (1):

$$\max|y|\leqq\frac{1}{2}K_1.$$

Hieraus folgt aber wegen

$$J_{\alpha+\beta}=\int_{-\infty}^{\infty}|y|^{\alpha+\beta}dx\leqq(\max|y|)^\beta\int_{-\infty}^{\infty}|y|^\alpha dx=(\max|y|)^\beta J_\alpha:$$

$$\left(\frac{J_{\alpha+\beta}}{J_\alpha}\right)^{\frac{1}{\beta}} \leqq \frac{1}{2} K_1;$$

damit ist (3) im Falle $p=1$ bewiesen. Zugleich sieht man, daß Gleichheit nur für $y\equiv 0$ besteht. Daß die Konstante $\frac{1}{2}$ genau ist, sieht man, wenn man etwa die Folge

$$y_n(x)=\begin{cases} 1 & \text{für } |x|\leqq 1 \\ 1-n(|x|-1) & \text{für } 1<|x|<1+\frac{1}{n} \\ 1 & \text{für } |x|\geqq 1+\frac{1}{n} \end{cases}$$

betrachtet.

Es sei nun $p>1$.

Wir wissen schon, daß aus der Endlichkeit von J_α und K_p folgt, daß $y(x)$ für $|x|\to\infty$ gegen 0 strebt, und, daß folglich $|y(x)|$ ihr Maximum annimmt. Da die Ungleichung (2) gegenüber einer Verschiebung $x\to x+\xi$ invariant ist, und da das Ersetzen von y durch ϱy das Multiplizieren beider Seiten mit $|\varrho|^{\alpha+\beta}$ bewirkt, darf man sich auf die Betrachtung solcher Funktionen $y(x)$ beschänken, für welche $y(0)=\max|y|=1$. Dann ist $|y(x)|^{\alpha+\beta}\leqq|y(x)|^\alpha$, also $J_{\alpha+\beta}\leqq J_\alpha$.

Es sei $f(s)=(s^\alpha-s^{\alpha+\beta})^{\frac{p-1}{p}}$ $(0\leqq s\leqq 1)$; man hat offenbar für $0<s<1$:

$$f'(s)=\frac{p-1}{p}\,\frac{\alpha s^\alpha-(\alpha+\beta)s^{\alpha+\beta}}{s[f(s)]^{\frac{1}{p-1}}}. \tag{11}$$

Es sei ferner $F(u)=\int_0^u f(s)\,ds$ $(0\leqq u\leqq 1)$.

Mit $y(x)$ ist auch $F(|y(x)|)$ eine totalstetige Funktion von x; man hat nur zu bemerken, daß (wegen $0\leqq f(s)\leqq 1$):

$$|F(|y(x_1)|)-F(|y(x_2)|)|=\left|\int_{|y(x_2)|}^{|y(x_1)|} f(s)\,ds\right|\leqq$$
$$\leqq\big||y(x_1)|-|y(x_2)|\big|\leqq|y(x_1)-y(x_2)|.$$

$F(|y(x)|)$ ist also die Integralfunktion ihrer Derivierten:

$$\frac{d}{dx}F(|y(x)|)=f(|y(x)|)\,[\operatorname{sgn} y(x)]\,y'(x).$$

Anwendung der Hölderschen Ungleichung ergibt:

$$
(12)\qquad \begin{aligned}
\int_{-\infty}^{\infty} f(|y|)\,|y'|\,dx &\leqq \Big(\int_{-\infty}^{\infty}[f(|y|)]^{\frac{p}{p-1}}dx\Big)^{\frac{p-1}{p}}\Big(\int_{-\infty}^{\infty}|y'|^p dx\Big)^{\frac{1}{p}} = \\
&= (J_\alpha - J_{\alpha+\beta})^{\frac{p-1}{p}}\,K_p^{\frac{1}{p}}.
\end{aligned}
$$

Da mit $y(x)$ auch $F(|y(x)|)$ gegen 0 strebt, wenn $|x|\to\infty$, erhält man weiter:

$$
(13)\qquad \begin{aligned}
\int_{-\infty}^{\infty} f(|y|)\,|y'|\,dx &\geqq \Big[-\int_0^{\infty} + \int_{-\infty}^{0}\Big](\operatorname{sgn} y) f(|y|) y'\,dx = \\
&= \Big[-\int_0^{\infty} + \int_{-\infty}^{0}\Big][F(|y(x)|)]'\,dx = 2F(|y(0)|) = 2F(1).
\end{aligned}
$$

Also hat man

$$2F(1) \leqq (J_\alpha - J_{\alpha+\beta})^{\frac{p-1}{p}}\,K_p^{\frac{1}{p}}.$$

Wenn man die Bezeichnung $V = \dfrac{J_{\alpha+\beta}}{J_\alpha}$ einführt, dann ist also

$$1 \leqq \frac{1}{2F(1)}\,(1-V)^{\frac{p-1}{p}}\,J_\alpha^{\frac{p-1}{p}}\,K_p^{\frac{1}{p}},$$

und folglich

$$V^{\frac{q}{\beta}} \leqq \frac{1}{2F(1)}\,V^{\frac{q}{\beta}}(1-V)^{\frac{p-1}{p}}\,J_\alpha^{\frac{p-1}{p}}\,K_p^{\frac{1}{p}}.$$

V ist offenbar eine Größe zwischen 0 und 1, daher

$$(14)\qquad V^{\frac{q}{\beta}}(1-V)^{\frac{p-1}{p}} \leqq \max_{0\leqq v\leqq 1} v^{\frac{q}{\beta}}(1-v)^{\frac{p-1}{p}} = M.$$

Also ist endlich

$$V^{\frac{q}{\beta}} \leqq \frac{M}{2F(1)}\,J_\alpha^{\frac{p-1}{p}}\,K_p^{\frac{1}{p}}.$$

Dies ist aber eben die Ungleichung (3). Denn es ist, wie eine leichte Rechnung ergibt, erstens

$$M = \frac{\left(\frac{q}{\beta}\right)^{\frac{q}{\beta}}\left(\frac{p-1}{p}\right)^{\frac{p-1}{p}}}{\left(\frac{q}{\beta}+\frac{p-1}{p}\right)^{\frac{q}{\beta}+\frac{p-1}{p}}},$$

zweitens

$$F(1)=\int_0^1 [s^\alpha(1-s^\beta)]^{\frac{p-1}{p}} ds = \frac{1}{\beta}\int_0^1 (1-u)^{\frac{p-1}{p}} u^{\frac{q}{\beta}-1} du =$$

$$=\frac{1}{\beta}\,\frac{\Gamma\left(\frac{q}{\beta}\right)\Gamma\left(1+\frac{p-1}{p}\right)}{\Gamma\left(1+\frac{q}{\beta}+\frac{p-1}{p}\right)}=\frac{1}{q}\,\frac{\Gamma\left(1+\frac{q}{\beta}\right)\Gamma\left(1+\frac{p-1}{p}\right)}{\Gamma\left(1+\frac{q}{\beta}+\frac{p-1}{p}\right)},$$

also

$$\frac{M}{2F(1)}=\frac{q}{2}H\left(\frac{q}{\beta},\frac{p-1}{p}\right).$$

Nun gehen wir zur Untersuchung des Bestehens des Gleichheitszeichens in (2) für $p>1$ über.

Damit in (2) das Gleichheitszeichen besteht, muß es gleichzeitig in (12), (13) und (14) bestehen. In (12) wird Gleichheit sein, wenn $|y'|^p$ ein konstantes Multiplum von $f^{\frac{p}{p-1}}(|y|)$ ist, d. h., wenn (fast überall)

$$(15) \qquad |y'|=\varrho f^{\frac{1}{p-1}}(|y|)$$

mit einer positiven Konstanten ϱ gilt. In (13) wird Gleichheit dann und nur dann bestehen, wenn fast überall, wo $f(|y|)\neq 0$ (also, wo $y\neq 0, \pm 1$), gilt:

$$(16) \qquad \begin{aligned} |y'| &= (\operatorname{sgn} y)\, y' \quad \text{für} \quad x<0,\\ |y'| &= -(\operatorname{sgn} y)\, y' \quad \text{für} \quad x>0. \end{aligned}$$

Endlich ist in (14) Gleichheit dann, wenn die Funktion $v^{\frac{q}{\beta}}(1-v)^{\frac{p-1}{p}}$ ihr Maximum gerade für $v=V$ erreicht, d. h., wenn

$$\frac{\frac{q}{\beta}}{V}-\frac{\frac{p-1}{p}}{1-V}=0,$$

oder, was damit gleichbedeutend ist, wenn

$$(17) \qquad J_{\alpha+\beta}\left(1+\frac{p-1}{p}(\alpha+\beta)\right)=J_\alpha\left(1+\frac{p-1}{p}\alpha\right).$$

Bedingung (16) ist offenbar damit gleichbedeutend, daß $(y^2)'$, d. h. $2yy'$ fast überall, wo $f(|y|)\neq 0$, für $x<0$ nichtnegativ, für $x>0$ aber nichtpositiv ausfällt. Wo aber $f(|y|)=0$, dort hat man wegen (15) fast überall: $y'=0$, also $(y^2)'=0$. Daraus folgt, daß $y^2(x)$ für $x<0$ monoton wächst, für $x>0$ aber monoton fällt.

Wegen $y(0) > 0$ ist dies dann und nur dann der Fall, wenn $y(0)$ überall nichtnegativ, für $x < 0$ monoton wachsend, für $x > 0$ aber monoton fallend ist. Für solche Funktionen nimmt (15) die Gestalt an:

$$(18)\quad y' = \varrho f^{\frac{1}{p-1}}(y) \quad \text{für} \quad x < 0, \quad -y' = \varrho f^{\frac{1}{p-1}}(y) \text{ für } x > 0.$$

Nehmen wir augenblicklich an, daß (18) eine zulässige Lösung $y(x)$ hat[3]), die überall nichtnegativ, für $x < 0$ wachsend, für $x > 0$ fallend ist. Es sei (σ, τ) das Intervall, wo $y(x)$ gleich 1 ist (es ist offenbar $\sigma \leqq 0 \leqq \tau$); ω_1 und $-\omega_2$ sollen die kleinste positive, bzw. die größte negative Zahl bedeuten, für die $y(x)$ verschwindet (falls es solche gibt, sonst sollen ω_1 das Symbol ∞, $-\omega_2$ das Symbol $-\infty$ bedeuten).

Mit Rückblick auf (11) haben wir dann

$$J_{\alpha+\beta} - J_\alpha + \frac{p-1}{p}\left[(\alpha+\beta) J_{\alpha+\beta} - \alpha J_\alpha\right] =$$

$$= \int_{-\infty}^{\infty} \left\{ y^{\alpha+\beta} - y^\alpha + \frac{p-1}{p}\left[(\alpha+\beta)\, y^{\alpha+\beta} - \alpha\, y^\alpha\right] \right\} dx = \int_\sigma^\tau + \int_{-\omega_2}^{\sigma} + \int_\tau^{\omega_1} =$$

$$= (\tau - \sigma)\frac{p-1}{p}\beta - \left[\int_{-\omega_2}^{\sigma} + \int_\tau^{\omega_1}\right]\left[f^{\frac{p}{p-1}}(y) + y f^{\frac{1}{p-1}}(y) f'(y)\right] dx.$$

Nun ist wegen (18)

$$\int_\tau^{\omega_1} = \int_1^0 \left[f^{\frac{p}{p-1}}(y) + y f^{\frac{1}{p-1}}(y) f'(y)\right] \frac{dx}{dy}\, dy = \int_0^1 [\ldots] \frac{dy}{\varrho f^{\frac{1}{p-1}}(y)} =$$

$$= \frac{1}{\varrho}\int_0^1 \left[f(y) + y f'(y)\right] dy = \frac{1}{\varrho}\left[y f(y)\right]_0^1 = 0;$$

ebenso erhält man, daß $\int_{-\omega_2}^{\sigma} = 0$. Hieraus folgt, daß (17) dann und nur dann erfüllt ist, wenn $(\tau - \sigma)\frac{p-1}{p}\beta = 0$, also wenn $\tau = \sigma = 0$.

Man hat also nur diejenigen Lösungen der Differentialgleichung (18) zu bestimmen, für die $y(0) = 1$ und $0 \leqq y(x) < 1$ für $x \neq 0$. Es existiert aber offenbar eine einzige solche Lösung, und

[3]) D. h., für die J_α und K_p endlich sind.

zwar $y = y_{p\alpha\beta}(|\varrho x|)$, wo $y_{p\alpha\beta}(t)$ die in § 1 definierte Funktion ist. Diese Funktion $y = y_{p\alpha\beta}(|\varrho x|)$ ist zulässig, d. h., die entsprechenden Integrale J_α und K_p sind endlich. Im Falle $\alpha < p$ ist das klar, da in diesem Falle $y_{p\alpha\beta}(t)$ außerhalb eines endlichen Intervalles verschwindet. Im Falle $\alpha \geqq p$ aber hat man

$$2J_\alpha = \int_0^\infty y^\alpha_{p\alpha\beta}(\varrho x)\,dx = \frac{1}{\varrho}\int_0^\infty y^\alpha_{p\alpha\beta}(t)\,dt = \frac{1}{\varrho}\int_1^0 u^\alpha \frac{dt}{du}\,du\,;$$

nun ist wegen (4): $\dfrac{dt}{du} = -[u^\alpha(1-u^\beta)]^{-\frac{1}{p}}$, also

$$2J_\alpha = \frac{1}{\varrho}\int_0^1 u^{\alpha-\frac{\alpha}{p}}(1-u^\beta)^{-\frac{1}{p}}\,du < \infty.\,^{4)}$$

Mit $y^\alpha_{p\alpha\beta}(|\varrho x|)$ hat $y^{\alpha+\beta}_{p\alpha\beta}(|\varrho x|)$ und also auch $y^\alpha_{p\alpha\beta}(|\varrho x|) - y^{\alpha+\beta}_{p\alpha\beta}(|\varrho x|)$ ein endliches Integral. Dann hat aber, wegen (15), auch $[y'_{p\alpha\beta}(|\varrho x|)]^p$ ein endliches Integral.

Wenn man die Bedingung $y(0) = \max|y| =$ wieder wegläßt, dann wird in (2) das Gleichheitszeichen für jene Funktionen gelten, die aus den Funktionen $y_{p\alpha\beta}(|\varrho x|)$ durch Verschiebung längs der x-Achse und Multiplizieren mit einer Konstanten hervorgehen.

(Eingegangen am 4. Dezember 1940.)

---

4) Es ist ja $\alpha - \dfrac{\alpha}{p} = \alpha\dfrac{p-1}{p} > 0$ und $-\dfrac{1}{p} > -1$.

Another proof of the mean ergodic theorem.

By FREDERICK RIESZ in Szeged.

Von Neumann's "mean ergodic theorem" will be given a new short proof in the following generalized form, established by the author some three years ago, together with other theorems of the same kind, concerning function spaces L^p.[1])

Theorem. *Let T be a bounded linear transformation in Hilbert space, for which $M_T \leqq 1$, i. e. $|Tf| \leqq |f|$ for any element f. Let f_1 be an element, let $f_n = T^{n-1} f_1$ and let φ_n be the arithmetic mean of $f_1, \ldots, f_n$. Then the sequence φ_n converges to a limit φ. More generally,*

$$\varphi_{m,n} = \frac{1}{n-m} \sum_{m+1}^{n} f_i \to \varphi \qquad (n-m \to \infty).$$

Proofs of the above theorem and of its extensions to more or less general spaces have been given on repeated occasions by several authors. In particular, it is with the proof given by GARRETT BIRKHOFF[2]) that our new argument has some features in common. Neither of them depends on weak compactness in spite of using minimal methods; in addition, both hold for any uniformly convex space. For sake of convenience, we write down the present argument for HILBERT space only; but it clearly may be read so as to embody the general case.

Birkhoff's starting-point is the infimum of the norms $|\varphi_n|$; instead, we consider the infimum μ of the norm $|g|$ on the whole

[1]) F. RIESZ, Some mean ergodic theorems, *Proceedings London Math. Society*, **13** (1938), pp. 274–278.

[2]) G. BIRKHOFF, The mean ergodic theorem, *Duke Math. Journal*, **5** (1939), pp. 19–20; *cf.* also L. ALAOGLU and G. BIRKHOFF, General ergodic theorems, *Annals of Math.*, **41** (1940), pp. 293–309.

convex set G of elements g of the type

$$(1) \qquad g=\sum_{1}^{\nu} c_i f_i \quad \left(\nu \text{ arbitrary, } c_i \geqq 0, \ \sum_{1}^{\nu} c_i = 1\right).$$

Clearly, the means φ_n and $\varphi_{m,n}$ belong to the set G.

In order to prove our theorem, we propose to show first that the sequence $\{\varphi_n\}$ and more generally, any sequence $\{\varphi_{m,n}\}$ with $n-m\to\infty$ is a minimizing sequence, *i. e.*, $|\varphi_n|\to\mu$ and $|\varphi_{m,n}|\to\mu$. As

$$|\varphi_{m,n}|=|T^m\varphi_{n-m}|\leqq|\varphi_{n-m}|,$$

the case of $\varphi_{m,n}$ in fact reduces to that of φ_n and all we have to show is that for any $\varepsilon>0$,

$$(2) \qquad |\varphi_N|<\mu+\varepsilon$$

for large N. To this purpose, let us take any element g of the type (1) with

$$(3) \qquad |g|<\mu+\frac{\varepsilon}{2}$$

and observe that the arithmetic mean ψ_N of $g, Tg, \ldots, T^{N-1}g$, when expressed in terms of f_1 and its transforms f_k, differs from φ_N, for $N>\nu-1$, but in its first and last terms, of the form $a_k f_k$ with $0\leqq a_k\leqq 1/N$, $2(\nu-1)$ in number, and so clearly

$$(4) \qquad |\psi_N-\varphi_N|\leqq\frac{2(\nu-1)}{N}|f_1|<\frac{\varepsilon}{2}$$

for large N. As, in addition, on account of the hypothesis $M_T\leqq 1$, (3) holds with $T^k g$ instead of g and thus for the mean ψ_n too, the inequality (2) readily follows from (4). Therefore, $\{\varphi_n\}$ and $\{\varphi_{m,n}\}$ are minimizing sequences.

Now, to finish our argument, we only have to recall the well known fact stating that for a convex set in HILBERT space, minimizing sequences are always convergent[3]). So, in particular, this is the case for the sequences $\{\varphi_{m,n}\}$ which proves the theorem.

(Received March 8, 1941)

---

[3]) See *e. g.* B. DE SZ. NAGY, On the set of positive functions in L_2, *Annals of Math.*, **39** (1938), pp. 1—13, in particular p. 5, footnote.

Bibliographie.

Maurice Lecat, Erreurs de Mathématiciens des origines à nos jours, XII + 167 pages, Bruxelles et Louvain, Castaigne et Desbarax, 1935.

Une liste étendue des erreurs, commises par des mathématiciens, à partir d'ARISTOTE, d'EUCLIDE et de SAINT AUGUSTIN jusqu'à ceux qui s'efforcent à démontrer le dernier théorème de FERMAT, avec des références bibliographiques concernant ces erreurs ainsi que leurs rélévations. Évidemment, une telle liste ne pourra nous être utile que si elle est complète relativement aux propositions erronées, figurant dans des périodiques qui leur pourriont conférer une autorité. On se convaincra par quelques épreuves qu'elle ne l'est pas. L. Kalmár.

Rudolf Rothe, Höhere Mathematik für Mathematiker, Physiker und Ingenieure, Teil IV: Übungsaufgaben mit Lösungen, Formelsammlung, unter Mitwirkung von OSKAR DEGOSANG, 3. Heft: Integralrechnung, 4. Heft: Unendliche Reihen, Vektorrechnung nebst Anwendungen, 5. Heft: Raumkurven und Flächen, Linienintegrale und mehrfache Integrale, 6. Heft: Gewöhnliche und partielle Differentialgleichungen nebst Anwendungen (Teubners math. Leitfäden, Band 35—38), III + III + 106 + III + III + 105 S., Leipzig und Berlin, B. G. Teubner, 1936—1937.

Von den vorliegenden Heften beziehen sich die Hefte 3 und 4 auf den Stoff des Teiles III des Werkes *Höhere Mathematik* von R. ROTHE. Hefte 5 und 6 umfassen Übungsmaterial aus dem Teile III. Eine Formelsammlung, die noch erscheint, wird den Abschluß des ganzen Werkes bilden.

Auch diese Hefte zeigen die schon lange erkannten Vorzüge der früher erschienenen übrigen Teile. Die Übungsbeispiele sind aus den verschiedensten Wissenszweigen sorgfältig ausgewählt und zusammengestellt. Der Leser gewinnt durch dieses Übungsmaterial nicht nur eine gute Übung im Rechnen, sondern auch wertvolle neue Erkenntnisse in der Mathematik und in ihren Anwendungen. Dank des vorzüglichen methodischen Geschickes der Verfasser kann diese Sammlung von Aufgaben allen Studierenden der Mathematik, Physik, Chemie und Technik, sogar auch den Dozenten der Mathematik, wärmstens empfohlen werden. Gy. v. Sz. N.

Andreas Speiser, Die Theorie der Gruppen von endlicher Ordnung (Grundlehren der math. Wissenschaften, Band V), dritte Auflage, X + 262 S., Berlin, J. Springer, 1937.

In der dritten Auflage des beliebten Lehrbuchs von SPEISER findet man nur wenig Veränderungen gegenüber der zweiten (s. hierüber das

Referat in *diesen Acta,* 4, S. 123). Inhaltliche Erweiterungen sind der Satz von HALL über die Automorphismengruppe von p-Gruppen, die Darstellung der symmetrischen Permutationsgruppen und ein neues Kapitel über die Theorie der algebraischen In- und Kovarianten. Dadurch brachte Verfasser das Werk seinem Nebenzweck, möglichst viel Ausblicke aus der Gruppentheorie in verschiedene Gebiete der Mathematik zu bieten, noch näher. Neu aufgenommen sind noch die schönen, einfachen Beweise von E. WITT für den Satz von FROBENIUS (*Berliner Berichte,* 1901) und den von WEDDERBURN über die Kommutativität endlicher Schiefkörper. Sonst wurde das Alte beibehalten, bis auf einige neue Literaturhinweise.

L. Rédei.

Paul B. Fischer, Arithmetik (Sammlung Göschen, 47), 152 S., Berlin, Walter de Gruyter, 1938.

Unter glücklicher Vereinigung der Knappheit eines Göschenbandes mit der Klarheit eines guten Lehrbuches entwickelt Verfasser die Lehre vom Rechnen mit natürlichen, mit ganzen, mit rationalen, mit reellen und schließlich mit komplexen Zahlen, einschließlich Kettenbrüche, Quaternionen, arithmetische und geometrische Reihen, Zinseszins- und Rentenrechnung und Kombinatorik. Die Darstellung ist anschaulich-genetisch; auch die historischen Gesichtspunkte werden hineinbezogen.

L. Kalmár.

American Mathematical Society Semicentennial Publications, Volume One: History, Volume Two: Addresses, XI + 262 + V + 315 pages, New York, American Mathematical Society, 1938.

In the autumn of 1888, the enthusiasm of three graduate students of Columbia University led to the founding of the New York Mathematical Society, which, within a few years, expanded into the American Mathematical Society. Now, on the threshold of the second half century of its steadily rising activity, it is with pride the Society can look back on its achievements during the first half, witnessed by more than a hundred volumes of its *Bulletin, Transactions* and *Colloquium Lectures.* To these, on the occasion of the Jubilee, the two present volumes have been added.

The first volume, Professor Archibald's work, is in its first part rather a biography of the Society as a whole, telling us about its childhood and its growing up, of its financial affairs, meetings, and publications. The larger part of the volume, with many pictures, is a memorial of the three chief secretaries and the twenty-four presidents, recording their lives and their publications; it may be considered as a reference book on prominent American mathematicians.

The second volume collects a review of nine invited addresses delivered at the Jubilee Meeting. Most of them are brief treatises surveying the recent contributions, in the States as well as abroad, to various special fields of mathematics; other refer to the whole fifty-years period and are concerned chiefly with research work done in America. Written by men who had a leading part in the work they are talking about, these papers will be of high value to the mathematical public.

Here follows the list of the papers: E. T. BELL, Fifty years of algebra in America, 1888 to 1938; J. F. RITT, Algebraic aspects of the theory of differential equations; NORBERT WIENER, The historical background of harmonic analysis; E. J. MCSHANE, Recent developments in the calculus of variations; T. Y. THOMAS, Recent trends in geometry; R. L. WILDER, The sphere in topology; G. C. EVANS, Dirichlet problems; J. L. SYNGE, Hydrodynamical stability; G. D. BIRKHOFF, Fifty years of American mathematics.

F. R.

D. Hilbert und P. Bernays, Grundlagen der Mathematik, zweiter Band (Grundlehren der math. Wissenschaften, Band L), XII + 498 S., Berlin, J. Springer, 1939.

Zur Zeit der Erscheinung des ersten Bandes war noch die Beweisbarkeit der Widerspruchsfreiheit der Arithmetik mit Hilfe der Beweistheorie nur eine starke, auch durch Gödels negativen Ergebnisse nicht erschütterbare Hoffnung Hilberts, die nicht jeder Kenner der Schwierigkeiten teilte. Inzwischen wurde diese Hoffnung durch Gentzens Beweis erfüllt. Es ist daher kein Wunder, daß man allgemein erwartet hat, der zweite Band des Hilbert-Bernaysschen Werkes werde sich hauptsächlich mit dem Gentzenschen Widerspruchsfreiheitsbeweis, sowie mit der Frage seiner Ausdehnbarkeit auf die Analysis befassen. Nun hat man sich in dieser Erwartung getäuscht: der Gentzensche Beweis wird zwar berücksichtigt im Sinne, daß — nach Darstellung der Gödelschen Sätze — jene Erweiterung des ursprünglich zu stark gespannten methodischen Anforderungen der Beweistheorie auseinandergesetzt wird, die erst den Gentzenschen Beweis trotz den Gödelschen Sätzen möglich gemacht hat; jedoch wird der Gentzensche Beweis selbst nur in seinen allgemeinsten Hauptgedanken skizziert.

Der Verzicht auf eine vollständige Darstellung des Gentzenschen Beweises hat wohl vor allem den Grund, daß derselbe in seiner ursprünglichen Formulierung noch nicht genügend durchsichtig dargestellt wurde und eine neue, vereinfachte Fassung erst während der Drucklegung des vorliegenden Bandes erscheinen ist. Außerdem verwendet GENTZEN auch in der neuen Fassung statt des im ersten Bande des vorliegenden Werkes entwickelten eine andere, von ihm in seiner Dissertation konstruierte Formalisierung der Logik, die zwar wesentliche heuristische Vorteile über jene besitzt, doch eine Neubearbeitung eines beträchtlichen Teiles des ersten Bandes erfordert hätte. Zwar läßt sich der Gentzensche Beweis, wie Referent — ebenfalls während der Drucklegung des vorliegenden Bandes — Herrn BERNAYS mitgeteilt hat, so umgestalten, daß man doch bei der ursprünglichen Formalisierung bleibt; diese Mitteilung wurde auch in einer Anmerkung während des Druckes, sowie in der Einleitung berücksichtigt; doch ist Referent auch noch jetzt einer Veröffentlichung des modifizierten Beweises schuldig. Auch der Ackermannsche Beweis der Widerspruchsfreiheit der Arithmetik, der sich nicht nur der Formalisierung, sondern auch den im vorliegenden Band dargelegten Hilbertschen Ansätzen unmittelbar anschließt, ist erst ein Jahr später erschienen als jener Band und war zur Zeit der Drucklegung desselben — wie aus der Einleitung hervorgeht — nicht fertig.

Ein tieferes Studium des Werkes zeigt aber klar, daß die aufgezählten Gründe nicht die entscheidendsten sind. Es würde nämlich auch dann nicht die Aufgabe des Werkes gewesen, den Gentzenschen Beweis darzustellen, wenn derselbe rechtzeitig und in unmittelbar in das Werk passender Form erschienen wäre — trotzdem, daß jener Beweis das zur Zeit erlangte tiefste positive Ergebnis der Beweistheorie liefert. Einfach darum nicht, weil das Werk wichtigere Aufgaben zu erfüllen hatte. Hilbert hatte nämlich seine grundlegenden Ideen und Ansätze zum Widerspruchsfreiheitsbeweis, insbesondere jene im Zusammenhang mit dem ε-Symbol („ausgezeichnetes Element derjenigen, welche") bloß andeutungsweise publiziert. Nun erhalten aber Hilberts diesbezügliche Untersuchungen eine Fülle von Ideen, aus denen die beweistheoretischen Forscher gewiß noch jahrelang — jedenfalls bis zur Erledigung der Widerspruchsfreiheitsfrage der Analysis — schöpfen können. Es war daher eine wissenschaftliche Verpflichtung der Verfasser gegenüber der Nachwelt, diese Untersuchungen in voller Ausführlichkeit darzustellen. Dieser Darstellung, die man als beweistheoretisches Testament Hilberts bezeichnen kann, ist fast die erste Hälfte des Buches gewidmet; erst dann folgt der bereits erwähnte Gödel-Gentzensche Ideenkreis und, in Supplementen, einige, zum Teil das Verständnis erleichternde, zum Teil aber gewisse Teile des Werkes ergänzende Ausführungen. Man kann nur bedauern, daß ein Teil des geistigen Schatzes Hilberts aus diesem Testament ausgeblieben ist; nach den Worten der Einleitung deshalb, da sie „teils vereinzelte Bemerkungen geblieben sind, teils noch keine hinlängliche Abklärung erfahren haben". Auch einen von HILBERT offenbar stark beeinflußten älteren Ansatz Ackermanns, die Zurückführung der Widerspruchsfreiheit der Analysis auf ein gewisses rein zahlentheoretisches Problem, die man nur aus mündlichen Mitteilungen kennt, entbehrt man nicht gern, da dessen Veröffentlichung vielleicht die Lösung der Widerspruchsfreiheitsfrage der Analysis fördern würde; doch liegt dieser geistige Schatz im Besitz Ackermanns.

In der Besprechung des ersten Bandes (*diese Acta*, **7**, S. 255) hat Referent gesagt, daß „man das Lesen des Buches nicht vor Abschluß des Paragraphen unterbrechen kann". Von dem vorliegenden Band gilt dies in erhöhtem Maße.

L. Kalmár.

Über die Verteilung der Primzahlen (I).

Von P. TURÁN in Budapest.

§ I.

In seinem Vortrag zu Cambridge in 1912 bezeichnete Professor LANDAU als eines der Hauptprobleme der Primzahltheorie den Beweis folgender Behauptung: wenn x eine genügend große, sonst aber beliebige Zahl bedeutet, dann gibt es zwischen x^2 und $(x+1)^2$ oder — was dasselbe bedeutet — zwischen x und $x+2\sqrt{x}+1$ immer Primzahlen. In diesem Aufsatz wollen wir uns mit dieser Frage beschäftigen.

Das Problem ist offenbar mit der Abschätzung von oben des Restgliedes der Primzahlformel verbunden. Wenn wir die berühmte Riemannsche Vermutung $\left(\text{d. h. } \zeta(s) \neq 0 \text{ für } \sigma > \frac{1}{2},\ s = \sigma + it\right)$ als bewiesen annehmen und wenn p die sukzessiven Primzahlen durchläuft, dann gilt für $x \to \infty$ nach H. VON KOCH[1])

$$\pi(x) = \sum_{p \leqq x} 1 = \int_2^x \frac{dt}{\log t} + O(\sqrt{x} \log x). \tag{1}$$

Aus (1) ergibt sich leicht die Existenz zweier numerischen Konstanten c_1 und c_2 so, daß für $x \geqq c_1$ zwischen x und $x + c_2\sqrt{x}\log^2 x$ immer Primzahlen existieren. Dieses Ergebnis wurde von H. CRAMÉR[2]) verbessert, wieder unter Annahme der Riemannschen Vermutung, indem er $\log^2 x$ durch $\log x$ ersetzen konnte. Dieser Weg wurde

---

[1]) H. VON KOCH, Sur la distribution des nombres premiers, *Acta Math.*, **24** (1901), p. 159—182.

[2]) H. CRAMÉR, Some theorems concerning prime numbers, *Arkiv för Math.*, **15** (1920), No. 5.

aber ziemlich hoffnungslos, als es klargemacht wurde, daß die Abschätzung des Restgliedes von oben in (1) mit der Wahrheit der Riemannschen Vermutung *untrennbar* verbunden ist; aus dieser Bemerkung erhellt sich, daß in der Riemannschen Vermutung im Wesen schon eine Behauptung vorliegt, welche in einer verschleierten Form die Verteilung der Primzahlen reguliert. LANDAU[3]) ist noch weiter gegangen, er hat bekanntlich direkt einen arithmetischen Zusammenhang zwischen ζ-Wurzeln und Primzahlen vermutet.

Einen sehr wichtigen Schritt machte HOHEISEL[4]) im Jahre 1930. Mittels einer sehr originellen Methode bewies er ohne Hypothesen, daß für alle hinreichend große x zwischen x und $x+x^{1-\frac{1}{33000}}$ immer Primzahlen existieren. Er entdeckte, daß man zur asymptotischen Abschätzung von $\pi(x+x^a)-\pi(x)$ $(a<1)$ kein Gebrauch davon zu machen hat, daß $\zeta(s)\neq 0$ in der Halbebene $\sigma>a$ ist, es genügt zu wissen I. daß $\zeta(s)\neq 0$ für ein Bereich $\sigma>1-D\frac{\log\log(|t|+3)}{\log(|t|+3)}$, $|t|>c_3$, wo D, c_3 und ferner $c_4,\dots$ numerische Konstanten bedeuten, II. daß keine der Halbebenen $\sigma>\sigma_0\left(\frac{1}{2}<\sigma_0<1\right)$ „zu viele" ζ-Wurzeln enthalten kann. Was I anbelangt, mußte er nur Littlewoods bekannte Abschätzung numerisch verfolgen, für II genügte ihm eine, etwas verschärfte Form des bekannten Satz von CARLSON[5]), nach welchem für die nichttrivialen ζ-Wurzeln, $\varrho=\sigma_\varrho+it_\varrho$,

$$N(\sigma_0, T)\equiv\sum_{\substack{\varrho\\ \sigma_\varrho\geqq\sigma_0\\ 0<t_\varrho\leqq T}} 1<c_4T^{4\sigma_0(1-\sigma_0)}\log^6 T. \tag{2}$$

Der Exponent $1-\frac{1}{33000}$ wurde zunächst durch die Vergrößerung von D in I von H. HEILBRONN[6]) zu $1-\frac{1}{250}$ verbessert; etwas

[3]) E. LANDAU, *Handbuch der Lehre von der Verteilung der Primzahlen* (Leipzig, 1909). Siehe besonders p. 367—368.

[4]) G. HOHEISEL, Primzahlprobleme in der Analysis, *Sitzungsberichte der Preußischen Akademie, phys. math. Klasse*, 1930, p. 580—588.

[5]) F. CARLSON, Über die Nullstellen der Dirichletschen Reihen und der Riemannschen ζ-Funktion, *Arkiv för Math.*, **15** (1920), No. 20.

[6]) H. HEILBRONN, Über den Primzahlsatz von Herrn Hoheisel, *Math. Zeitschrift*, **36** (1933), p. 394—423.

später bewies TCHUDAKOFF[7]) mittels der Winogradoffschen Methode, daß D beliebig groß gewählt werden kann, woraus, wie schon HOHEISEL angibt, folgt, daß zwischen x und $x+x^{\frac{3}{4}+\varepsilon}$ immer Primzahlen existieren, wenn nur $x > c_5 = c_5(\varepsilon)$. Eine weitere wesentliche Verkleinerung des Exponenten $\frac{3}{4}+\varepsilon$ gelang INGHAM[8]), indem er entdeckte, daß eine Verbesserung des Exponenten $4\sigma_0(1-\sigma_0)$ in (2) eine Verkleinerung des Exponenten $\frac{3}{4}+\varepsilon$ nach sich zieht. Aus Hoheisels Beweis arbeitete er (Tchudakoffs Resultat wieder benützend) den folgenden Satz aus: wenn für b und B die Ungleichung

$$N(\sigma_0, T) < c_6 T^{b(1-\sigma_0)} \log^B T \tag{3}$$

gleichmäßig für $\frac{1}{2} \leqq \sigma_0 \leqq 1$ erfüllt ist, dann ist für jedes $\vartheta > \frac{1}{b}$

$$\lim_{x\to\infty} [\pi(x+x^\vartheta) - \pi(x)] \frac{\log x}{x^\vartheta} = 1. \tag{4}$$

Wie aus dem wohlbekannten Riemann—Mangoldtschen Satze $N\left(\frac{1}{2}, T\right) \sim c_7 T \log T$ folgt, kann b nicht kleiner sein als 2, d. h., (4) ist auf diesem Wege *grundsätzlich* nur für $\vartheta \geqq \frac{1}{2}$ beweisbar. INGHAM beweis (3) mit $b=\frac{8}{3}$ und so (4) mit $\vartheta > \frac{5}{8}$. Für $\vartheta < \frac{1}{2}$ sind mir nur Cramérs neuere Ergebnisse bekannt[9]), von denen typisch ist der folgende Satz: unter Annahme der Riemannschen Vermutung und $0 < \alpha < \frac{1}{2}$ gilt

$$\sum_{\substack{p_n \leqq x \\ p_{n+1} - p_n \geqq p_n^\alpha}} 1 < c_8 x^{1-2\alpha} \log^2 x.$$

Hier bedeuten $2 = p_1 < p_2 < \ldots$ die sukzessiven Primzahlen.

---

[7]) N. G. TCHUDAKOFF, On zeros of Dirichlets L-functions, *Recueil math. (Moscou)*, 1 (43) (1936), p. 591—602.

[8]) A. E. INGHAM, On the difference between consecutive primes, *Quarterly Journal*, Oxford Series, 8 (1937), p. 255—266.

[9]) H. CRAMÉR, On the order of the magnitude of the difference between consecutive prime-numbers, *Acta Aritmetica*, 1 (1936), p. 23—46.

INGHAM beweis allgemeiner (3) mit $b=2+4c$, $B=5$, wo c die untere Grenze derjenigen Exponenten β bedeutet, für welche $\zeta\left(\frac{1}{2}+it\right)=O(t^{\beta})$ gilt. Wenn also $c=0$, d. h. wenn die Lindelöfsche Vermutung wahr ist, dann bewies INGHAM wesentlich den oben erwähnten Koch—Cramérschen Satz, wo er aber statt der Riemannschen Vermutung die schwächere[10]) Lindelöfsche benötigte. Auch diese ist aber gewissermaßen mit den Primzahlen innig verbunden, da aus (1) die Riemannsche Vermutung und nach[10]) die Lindelöfsche folgt.

In § II werde ich die Ungleichung

(5) $$N(a, T)\leqq T^{2(1-a)}\exp(13\log^{0{,}18}T)$$

gleichmäßig für $\frac{1}{2}\leqq a\leqq 1$, $T\geqq c_9$ beweisen, vorläufig unter Annahme einer anderen Vermutung, welche aber meines Erachtens unabhängig von den Primzahlen ist. Diese Annahme wird enger, wenn wir (5) gleichmäßig für $\frac{1}{2}\leqq a\leqq 1$ beweisen wollen, als wenn der Beweis nur für $\frac{1}{2}\leqq a\leqq 1-\varepsilon$ geführt wird. Für letzteres genügt die

Hypothese **a.** *Es sei $\omega(n)$ eine positive, für $n\to\infty$ monoton ins Unendlich strebende Funktion von n, sonst beliebig. Für $n\geqq c_9$ gibt es ein $M=M(n)$, das die Ungleichungen*

(6a) $$n\omega(n)\leqq M(n)\leqq\frac{n^2}{\omega(n)^2}$$

erfüllt und es gilt

(6b) $$\max_{M\left(1-\frac{1}{\omega(n)}\right)\leqq\nu\leqq M}|z_1^{\nu}+z_2^{\nu}+\ldots+z_n^{\nu}|>\exp\left(-\frac{n^2}{M\omega(n)}\right)$$

für alle Wertsysteme $(z_1, z_2, \ldots, z_n)$, welche die Bedingungen

(6c) $$z_1=1,\ 1\geqq|z_\nu|\geqq 1-\frac{n^2}{2M^2},\quad (\nu=2, 3, \ldots, n)$$

genügen.

---

[10]) J. E. LITTLEWOOD, Quelques consequences de l'hypothèse que la fonction $\zeta(s)$ de Riemann n'a pas de zeros dans le demi plan $\sigma>\frac{1}{2}$. *Comptes Rendus Paris,* **154** (1912), p. 263—265.

Den wesentlichen Inhalt dieser Hypothese kann man offenbar so ausdrücken, daß „zu viele" konsekutive Potenzsummen von n komplexen Zahlen nicht gleichzeitig „zu klein" sein können. Übrigens zeigt eine heuristische Überlegung, daß diese Hypothese wahrscheinlich für *alle* $M(n)$, welche (6a) genügen, *alle* Wertsysteme $(z_1, z_2, \ldots, z_n)$, welche nur die schwächeren Bedingungen $z_1 = 1$, $|z_\nu| \leqq 1$ $(\nu = 2, 3, \ldots, n)$ erfüllen, gültig ist, sogar mit $\exp(-\log^3 n)$ statt $\exp\left(-\frac{n^2}{M\omega(n)}\right)$. In § III werden wir zwei Resultate in der Richtung der Hypothese entwickeln; der erste ergibt, daß, wenn nur die Bedingungen $M \geqq 2n$, $|z_\nu| \geqq 1 - \frac{n}{M}$ $(\nu = 1, 2, \ldots, n)$ erfüllt sind, dann gilt

$$(7) \qquad \max_{M-n+1 \leqq \nu \leqq M} |z_1^\nu + \ldots + z_n^\nu| \geqq \left(\frac{n}{200 M}\right)^n.$$

Zur Orientierung bemerken wir, daß, wenn wir $n\omega(n)^2$ als $M(n)$ wählen, dann gibt (7) die Abschätzung

$$(8a) \qquad \max_{M-n+1 \leqq \nu \leqq M} |z_1^\nu + \ldots + z_n^\nu| \geqq \exp\{-n \log(200\,\omega(n)^2)\}.$$

während (6b) die Ungleichung

$$(8b) \qquad \max_{M-n\omega(n) \leqq \nu \leqq M} |z_1^\nu + \ldots + z_n^\nu| \geqq \exp\left\{-\frac{n}{\omega(n)^3}\right\}$$

fordert. Wenn $\omega(n)$ „langsam" ins Unendlich strebt, so ist die rechte Seite von (8a) „fast" die gewünschte, trotzdem, daß in der bewiesenen Ungleichung (8a) das Intervall für ν viel enger ist, als in (8b). Das zweite Resultat liegt in der entgegengesetzten Richtung; ich gebe, eine Bemerkung von Herrn P. Stein folgend, ein Wertsystem $(z_1, z_2, \ldots, z_n)$ mit $z_1 = 1$, $|z_2| = |z_3| = \ldots = |z_n| = 1$ an, für welches

$$\max_{M\left(1-\frac{1}{\omega(n)}\right) \leqq \nu \leqq M} |z_1^\nu + z_2^\nu + \ldots + z_n^\nu| < \exp\left(-\frac{5}{6}\log n \,\log\omega(n)\right)$$

bei jeder Wahl von M gilt.

Wenn wir (5) gleichmäßig für $\frac{1}{2} \leqq a \leqq 1$ beweisen wollen, benötigen wir eine schärfere Hypothese, welche auch ein gewissermaßen willkürliches $M = M(n)$ enthält; da diese Hypothese wahrscheinlich für jede erlaubte Wahl von M und sogar für $z_1 = 1$,

$|z_2| \leqq 1,\ |z_3| \leqq 1, \ldots, |z_n| \leqq 1$ wahr ist, einfachkeitshalber benützen wir hier sie mit $M = n^{3/2}$. Dann lautet unsere Hypothese, wie folgt:

Hypothese b. *Für* $n \geqq c_{10}$ *und* $z_1 = 1,\ |z_2| \leqq 1, \ldots, |z_n| \leqq 1$ *gilt*

$$\max_{n^{3/2}(1-n^{-0{\cdot}42}) \leqq \nu \leqq n^{3/2}} |z_1^\nu + z_2^\nu + \ldots + z_n^\nu| > \exp(-n^{0{\cdot}09}).$$

Aus (5) leiten wir, die Beweisführung von HOHEISEL—INGHAM folgend, für $x \to \infty$

$$\pi(x + \sqrt{x} \exp(\log^{0{\cdot}996} x)) - \pi(x) \sim \frac{\sqrt{x}}{\log x} \exp(\log^{0{\cdot}996} x) \tag{9}$$

ab. So gelangen wir im Wesen zu dem Koch—Inghamschen Satze, aber unter Annahme einer *von Primzahlen unabhängigen* Vermutung. Hierbei sei noch bemerkt, daß ich für den Inghamschen Satz einen neuen Beweis gefunden habe — natürlich ohne irgendwelche Vermutungen — der mit Kochs Beweis einige Ähnlichkeit zeigt und der die „verkürzte" Primzahlformel von RIEMANN—MANGOLDT—LANDAU nicht benützt. Von der Darstellung dieses Beweises wollen wir diesmal absehen.

Der Ausgangspunkt dieser Untersuchungen war der folgende heuristische Beweis der Riemannschen Vermutung. Es sei $s_0 = \sigma_0 + it_0$ mit $\sigma_0 > 1,\ t_0^2 > 2\sigma_0 - 1$, so daß s_0 von dem Punkte $s = 1$ ferner liegt, als von der imaginären Achse. Nach dem Satze von CAUCHY—HADAMARD hat $\zeta(s)$ keine Wurzel im Kreise

$$|s - s_0|^{-1} = \overline{\lim_{\nu \to \infty}} \sqrt[\nu]{\frac{1}{\nu!} \left| \frac{\zeta'}{\zeta}(s)^{(\nu)} \right|_{s=s_0}} = R^{-1}.$$

Wenn wir heuristisch $\left| \frac{\zeta'}{\zeta}(s)^{(\nu)} \right|_{s=s_0}$ durch den längst der ganzen Linie $\sigma = \sigma_0$ genommenen quadratischen Mittelwert ersetzen, gelangen wir zu[11])

$$R^{-1} = \overline{\lim_{\nu \to \infty}} \sqrt[2\nu]{\frac{1}{\nu!^2} \sum_{k=1}^{\infty} \frac{\Lambda(k)^2 \log^{2\nu} k}{k^{2\sigma_0}}} \leqq \overline{\lim_{\nu \to \infty}} \sqrt[2\nu]{\frac{1}{\nu!^2} \sum_{k=1}^{\infty} \frac{\log^{2\nu+2} k}{k^{2\sigma_0}}}.$$

Wie man aber leicht sieht, ist der rechtsstehende $\overline{\lim}$ (sogar lim) genau $\left(\sigma_0 - \frac{1}{2}\right)^{-1}$; so ergibt sich, daß, wenn s_0 die Gerade $\sigma = \sigma_0$ durchläuft, $\zeta(s)$ in dem Bereiche $\sigma > \frac{1}{2}, |t| > \sqrt{2\sigma_0 - 1}$ keine Wur-

[11]) $\Lambda(k)$ bedeutet, wie üblich, $\log p$ für $k = p^\alpha$ und sonst 0.

zel hat. Vorliegender Aufsatz enthält eine erste Verbesserung dieser heuristischen Methode; ich hoffe bald auch Verbesserungen anderer Art mitteilen zu können. Zwecks Abschätzung von $N(a, T)$ nehmen wir jetzt die vorherigen ν und σ_0 als geeignete Funktionen von T an, für welche wir anfangs nur $\lim_{T\to\infty} \nu = \lim_{T\to\infty} \sigma_0 = +\infty$ voraussetzen; mit diesen ν und σ_0 ergibt die obere Abschätzung des Integrals

$$J = \int\limits_{\substack{T/2\\ \sigma=\sigma_0}}^{T} \left| \frac{\zeta'}{\zeta}(s)^{(\nu)} \right|^2 dt,$$

daß eine Ungleichung von der Form

$$\left| \frac{\zeta'}{\zeta}(s)^{(\nu)} \right| \leqq T^{a-1} \sqrt{J} \tag{10}$$

„sehr oft" für $T/2 \leqq t \leqq T$, $\sigma = \sigma_0$ gilt, d. h. überall in diesem Intervall höchstens mit Ausnahme einer Intervallmenge $\mathfrak{A} = \mathfrak{A}(\nu, a)$, deren Gesamtlänge $m(\mathfrak{A})$ die Ungleichung $m(\mathfrak{A}) < c_{11}(\varepsilon)\, T^{2(1-a)+\varepsilon}$ genügt. Diese Menge $\mathfrak{A}$ hängt selbst von ν ab; wenn wir aber den Wert von ν auf ein Intervall kürzer als T^{ε} beschränken, dann ist für $\sum\limits_{\nu} \mathfrak{A}(\nu, a) = \mathfrak{A}^{+}$ offenbar $m(\mathfrak{A}^{+}) < c_{12}(\varepsilon)\, T^{2(1-a)+2\varepsilon}$, d. h. (10) ist im $\sigma = \sigma_0$, $T/2 \leqq t \leqq T$ „sehr oft" *für beliebige zugelassene ν gleichzeitig* befriedigt. Aus der Riemann—Hadamardschen Formel

$$\begin{aligned} \frac{\zeta'}{\zeta}(s) = b - \frac{1}{s-1} + \frac{1}{s+2} + \\ + \sum_{n=1}^{\infty}\left(\frac{1}{s+2n+2} - \frac{1}{2n}\right) + \sum_{\varrho}\left(\frac{1}{s-\varrho} + \frac{1}{\varrho}\right) \end{aligned} \tag{11}$$

ausgehend (b Weltkonstante) können wir $\left| \frac{\zeta'}{\zeta}(s)^{(\nu)} \right|$ auf der Komplementärmenge $\mathfrak{B}$ von $\mathfrak{A}^{+}$ im $T/2 \leqq t \leqq T$ *nach einer geeigneter Wahl von ν* von unten abschätzen; diese beiden Abschätzungen ergeben, daß die Ungleichung $\sigma_\varrho \leqq a$ im Intervalle $T/2 \leqq t \leqq T$ „sehr oft" befriedigt ist.

Das Interessante in den oben skizzierten Methoden ist die Tatsache, daß ich, im Gegensatz zu den klassischen Methoden, immer mit Funktionswerten operiere, welche „fern von dem kritischen Streifen" angenommen werden.

Wir wenden folgende bekannte Sätze aus der Theorie der ζ-Funktion an. Es gibt eine positive, absolut konstante ganze Zahl $A \geqq 4$ mit den Eigenschaften:

(12) für $|t| \geqq A$ ist[12]
$$\sum_{\substack{\varrho \\ t \leqq t_\varrho \leqq t+\frac{1}{2}}} 1 > 0,$$

(13)
$$\sum_{-A \leqq t_\varrho \leqq A} 1 < \frac{A^2}{2},$$

(14) für $t \geqq A$ gilt[13]
$$\left| \sum_{0 \leqq t_\varrho \leqq t} 1 - \frac{t}{2\pi} \log \frac{t}{2\pi} + \frac{t}{2\pi} \right| < \frac{1}{2} \log t$$

und

(15)
$$\sum_{\substack{\varrho \\ t \leqq t_\varrho \leqq t+1}} 1 < \min(t-1, \log t).$$

Ferner erwähne ich noch drei einfache Ungleichungen, die in folgendem öfter benötigt werden: Wenn n eine positive ganze Zahl bedeutet, dann gilt

(16a)
$$\left(\frac{n}{e}\right)^n \leqq \frac{1}{e}\, n!,$$

(16b)
$$(n+1)!^2 \leqq \frac{(2n+3)!}{2^{2n+2}},$$

(16c)
$$(2n+2)! \leqq 4^{n+1} (n+1)!^2.$$

§ II.

Lemma I. *Es sei $b=1$ oder 2, $K \geqq 1$, $r \geqq 2$. Dann ist*
$$J_1 \equiv \sum_{n=1}^{\infty} \frac{\log^K (bn)}{n^r} < 2\, \frac{b^r K!}{(r-1)^K}.$$

Beweis: Da
$$\frac{d}{dy}\left(y^{-r} \log^K (by)\right) = y^{-r-1} \log^{K-1}(by)\,(K - r\log(by))$$

für $y > 1$ nur dann verschwindet, wenn $y = \frac{1}{b} e^{K/r}$, so haben wir nach (16a)

---

[12]) J. E. Littlewood, Two notes on the Riemann ζ-function, *Proceedings Cambridge Phil. Society,* **22** (1924), p. 234—242.

[13]) R. Backlund, Über die Nullstellen der Riemannschen Zetafunktion, *Acta Math.,* **41** (1928), p. 345—375.

$$J_1 < 2b^r\left(\frac{K}{er}\right)^K + \int_1^\infty \frac{\log^K(by)}{y^r}\,dy =$$

$$= 2b^r\left(\frac{K}{er}\right)^K + b^{r-1}\int_0^\infty e^{-(r-1)y} y^K dy = 2b^r\left(\frac{K}{er}\right)^K + \frac{b^{r-1} K!}{(r-1)^{K+1}} <$$

$$< 2b^r \frac{1}{e}\,\frac{K!}{r^K} + \frac{b^r K!}{(r-1)^K} < 2b^r \frac{K!}{(r-1)^K}.$$

Lemma II. *Für* $\nu > 0$, $\sigma_0 > 0$, $\frac{\sigma_0}{\nu+1} < 1$ *und* $1 \leqq n \leqq \frac{1}{2} e^{\frac{\nu+1}{\sigma_0}}$ *gilt*

$$\log^{\nu+1} n \leqq \frac{1}{2^{\sigma_0}} \log^{\nu+1}(2n).$$

Beweis:

$$\log n \leqq \frac{\nu+1}{\sigma_0} - \log 2 = \frac{(\nu+1) - \sigma_0 \log 2}{\sigma_0},$$

d. h.

$$1 + \frac{\log 2}{\log n} \geqq 1 + \frac{\sigma_0 \log 2}{(\nu+1) - \sigma_0 \log 2} = \frac{1}{1 - \frac{\sigma_0}{\nu+1}\log 2} \geqq e^{\frac{\sigma_0}{\nu+1}\log 2} = 2^{\frac{\sigma_0}{\nu+1}}.$$

Daraus folgt aber

$$\left(\frac{\log(2n)}{\log n}\right)^{\nu+1} \geqq 2^{\sigma_0}, \qquad \text{w. z. b. w.}$$

Wir haben

$$N(a, T) < T^{2(1-a)} \exp(13 \log^{0{\cdot}18} T)$$

für $\frac{1}{2} \leqq a \leqq 1$ und für $T > C$ zu beweisen; C bedeutet hier und später absolute, nicht notwendig dieselben Konstanten (also hier ist C speziell von a und T unabhängig). Es ist mehr bequemlich a durch $1 - \frac{\alpha}{2}$ zu ersetzen; dann ist unsere Behauptung

$$N\left(1 - \frac{\alpha}{2}, T\right) < T^\alpha \exp(13 \log^{0{\cdot}18} T) \tag{17}$$

für $T > C$ und $0 \leqq \alpha \leqq 1$. Da TITCHMARCH[14]) mittels der Wino-

---

[14]) E. C. TITCHMARCH, On $\zeta(s)$ and $\pi(x)$, *Quarterly Journal*, Oxford Series, 9 (1938), p. 97—108.

gradoffschen Methode

$$N(1-2\log^{-0{\cdot}81} T,\, T)=0$$

für $T>C$ bewiesen hat, haben wir (17) nur für

$$(18)\qquad T>C,\quad 1\geqq\alpha\geqq 2\log^{-0{\cdot}81} T$$

zu beweisen.

Wir betrachten zuerst

$$(19)\qquad N^{+}\left(1-\frac{\alpha}{2},\, T\right)=N\left(1-\frac{\alpha}{2},\, T\right)-N\left(1-\frac{\alpha}{2},\, \frac{T}{2}\right).$$

Es sei in folgendem

$$(20)\qquad \sigma_0=2^6\,(1+\log^{-0{\cdot}84} T)\log^2 T=B\log^2 T$$

und die ganze Zahl $\nu=\nu(T)$ beschränken wir in diesem Moment nur durch

$$(21)\qquad 2B\log^2 T\leqq\nu\leqq B\log^3 T.$$

Dann ist für $T>C$ offenbar

$$(22)\qquad \begin{gathered}\sigma_0>100,\ \frac{\sigma_0}{\nu+1}<1,\\ \left(1+\frac{1}{2\sigma_0-2}\right)^{2\nu+2}<e^{\frac{\nu+1}{\sigma_0-1}}<2T.\end{gathered}$$

Lemma III. *Mit den oben definierten σ_0 und ν gilt für $T>C$*

$$J_2\equiv\int\limits_{\substack{T/2\\ \sigma=\sigma_0}}^{T}\left|\frac{\zeta'}{\zeta}(s)^{(\nu)}\right|^2 dt<109\,T\log T\,\frac{(2\nu+3)!}{(2\sigma_0-1)^{2\nu+2}}.$$

Beweis: Aus der bekannten, für $\sigma>1$ gültigen Identität

$$\sum\frac{\Lambda(n)}{n^s}=-\frac{\zeta'}{\zeta}(s)$$

folgt:

$$(23)\qquad \sum\frac{\Lambda(n)\log^{\nu} n}{n^s}=(-1)^{\nu+1}\frac{\zeta'}{\zeta}(s)^{(\nu)}.$$

Also ist, wenn wir im Beweis des Lemma III σ_0 zur Abkürzung durch σ ersetzen,

$$J_2=\left|\frac{T}{2}\sum_n\frac{\Lambda(n)^2\log^{2\nu} n}{n^{2\sigma}}+\right.$$

$$\left.+\sum\sum_{m\neq n}\frac{\Lambda(m)\Lambda(n)\log^{\nu} m\log^{\nu} n}{(mn)^{\sigma}}\int\limits_{T/2}^{T}\left(\frac{m}{n}\right)^{it} dt\right|<$$

$$< \frac{T}{2}\sum_{n=2}^{\infty}\frac{\log^{2\nu+2} n}{n^{2\sigma}} + 4\sum_{n<m}\sum \frac{\log^{\nu+1} m \log^{\nu+1} n}{(mn)^{\sigma}\log\frac{m}{n}} =$$

$$= \frac{T}{2}\sum_{n=2}^{\infty}\frac{\log^{2\nu+2} n}{n^{2\sigma}} + 4\sum_{m\geqq 2n}\sum \frac{\log^{\nu+1} m \log^{\nu+1} n}{(mn)^{\sigma}\log\frac{m}{n}} +$$

$$+ 4 \sum_{n<m<2n}\sum \frac{\log^{\nu+1} m \log^{\nu+1} n}{(mn)^{\sigma}\log\frac{m}{n}} <$$

$$< \frac{T}{2}\sum_{n=2}^{\infty}\frac{\log^{2\nu+2} n}{n^{2\sigma}} + \frac{4}{\log 2}\left(\sum_{n=2}^{\infty}\frac{\log^{\nu+1} n}{n^{\sigma}}\right)^2 +$$

$$+ \frac{4}{\log 2}\sum_{n=2}^{\infty}\frac{\log^{\nu+1} n}{n^{\sigma-1}}\sum_{n<m<2n}\frac{\log^{\nu+1} m}{m^{\sigma}(m-n)},$$

da für $0\leqq y\leqq 1$ $\log(1+y)\geqq y\log 2$. Für die ersten zwei Glieder können wir Lemma I mit $K=2\nu+2$, $r=2\sigma$, $b=1$ bzw. mit $K=\nu+1$, $r=\sigma$, $b=1$ anwenden und so gelangen wir zu

$$|J_2| < \frac{(2\nu+2)!}{(2\sigma-1)^{2\nu+2}}T + \frac{16}{\log 2}\frac{(\nu+1)!^2}{(\sigma-1)^{2\nu+2}} +$$

$$\tag{24} + \frac{4}{\log 2}\sum_{n\leqq\frac{1}{2}e^{(\nu+1)/\sigma}}\frac{\log^{\nu+1} n}{n^{\sigma-1}}\left(\sum_{n<m<2n}\frac{\log^{\nu+1} m}{m^{\sigma}(m-n)}\right) +$$
$$+ \frac{4}{\log 2}\sum_{\frac{1}{2}e^{(\nu+1)/\sigma}<n\leqq e^{(\nu+1)/\sigma}}\frac{\log^{\nu+1} n}{n^{\sigma-1}}\left(\sum_{n<m<2n}\frac{\log^{\nu+1} m}{m^{\sigma}(m-n)}\right) +$$
$$+ \frac{4}{\log 2}\sum_{n>e^{(\nu+1)/\sigma}}\frac{\log^{\nu+1} n}{n^{\sigma-1}}\left(\sum_{n<m<2n}\frac{\log^{\nu+1} m}{m^{\sigma}(m-n)}\right).$$

Die letzten drei Summen seien mit J_3, J_4 und J_5 bezeichnet und es sei zuerst J_5 betrachtet. Da die Funktion $\frac{\log^{\nu+1} x}{x^{\sigma}}$ von $x=1$ bis $x=e^{(\nu+1)/\sigma}$ monoton zunimmt und für $x>e^{(\nu+1)/\sigma}$ monoton abnimmt, so ist

$$|J_5| < \frac{4}{\log 2}\sum_{n>e^{(\nu+1)/\sigma}}\frac{\log^{\nu+1} n}{n^{\sigma-1}}\frac{\log^{\nu+1} n}{n^{\sigma}}\sum_{n<m<2n}\frac{1}{m-n} <$$

$$< \frac{8}{\log 2}\sum_{n>e^{(\nu+1)/\sigma}}\frac{\log^{2\nu+3} n}{n^{2\sigma-1}} < \frac{8}{\log 2}\sum_{n=2}^{\infty}\frac{\log^{2\nu+3} n}{n^{2\sigma-1}}$$

und nach Lemma I (mit $b=1$, $K=2\nu+3$, $r=2\sigma-1$)

$$\tag{25a} |J_5| < \frac{16}{\log 2}\frac{(2\nu+3)!}{(2\sigma-2)^{2\nu+3}}.$$

Nun betrachten wir J_3. Da $2n \leqq e^{(\nu+1)/\sigma}$ gilt, gelangen wir nach der obigen Bemerkung über den Verlauf der Funktion $\dfrac{\log^{\nu+1} x}{x^\sigma}$ zu

$$|J_3| < \frac{4}{\log 2} \sum_{2 \leqq n \leqq \frac{1}{2} e^{(\nu+1)/\sigma}} \frac{\log^{\nu+1} n}{n^{\sigma-1}} \frac{\log^{\nu+1}(2n)}{(2n)^\sigma} \sum_{n<m<2n} \frac{1}{m-n}$$

und mit Rücksicht auf Lemma II

$$|J_3| < \frac{12}{\log 2} \frac{1}{2^{2\sigma}} \sum_{n=1}^{\infty} \frac{\log^{2\nu+3}(2n)}{n^{2\sigma-1}}.$$

Lemma I ergibt also (mit $b=2$, $K=2\nu+3$, $r=2\sigma-1$)

(25b) $$|J_3| < \frac{12}{\log 2} \frac{(2\nu+3)!}{(2\sigma-2)^{2\nu+3}}.$$

Wir schätzen endlich J_4 ab. Da für $x>1$ die Ungleichung $\dfrac{\log^{\nu+1} x}{x^\sigma} \leqq \left(\dfrac{\nu+1}{e\sigma}\right)^{\nu+1}$ gilt, so folgt mit Rücksicht auf (20) und (21) für $T>C$

$$|J_4| < \frac{4}{\log 2} \sum_{\frac{1}{2} e^{(\nu+1)/\sigma} \leqq n \leqq e^{(\nu+1)/\sigma}} \frac{\log^{\nu+1} n}{n^{\sigma-1}} \left(\frac{\nu+1}{e\sigma}\right)^{\nu+1} \left(\sum_{n<m<2n} \frac{1}{m-n}\right) <$$

$$< \frac{12}{\log 2} \left(\frac{\nu+1}{e\sigma}\right)^{\nu+1} \sum_{\frac{1}{2} e^{(\nu+1)/\sigma} \leqq n \leqq e^{(\nu+1)/\sigma}} \frac{\log^{\nu+2} n}{n^{\sigma-1}} <$$

$$< \frac{12}{\log 2} \left(\frac{\nu+1}{e\sigma}\right)^{\nu+1} \frac{1}{2} e^{\frac{\nu+1}{\sigma}} \left(\frac{\nu+2}{e(\sigma-1)}\right)^{\nu+2} <$$

$$< \frac{7}{\log 2} \left(\frac{\nu+1}{e}\right)^{2\nu+2} \left(\frac{e^{1/\sigma}}{\sigma(\sigma-1)}\right)^{\nu+1} \log T$$

und nach (16a) und (16b)

$$|J_4| < \frac{7}{e^2 \log 2} (\nu+1)!^2 \left(\frac{e^{1/\sigma}}{\sigma(\sigma-1)}\right)^{\nu+1} \log T <$$

$$< \frac{7}{e^2 \log 2} (2\nu+3)! \left(\frac{e^{1/\sigma}}{4\sigma(\sigma-1)}\right)^{\nu+1} \log T.$$

Da aber $(2\sigma-2)e^{1/\sigma} < (2\sigma-2)\dfrac{1}{1-\frac{1}{\sigma}} = 2\sigma$ gilt, ist

(25c) $$|J_4| < \frac{7}{e \log 2} \log T \frac{(2\nu+3)!}{(2\sigma-2)^{2\nu+2}}.$$

Aus (25a), (25b), (25c), (24) und (22) gewinnen wir für $T>C$

die Abschätzung

$$|J_2| < T\frac{(2\nu+2)!}{(2\sigma-1)^{2\nu+2}} + \frac{28}{\log 2}\,\frac{(2\nu+3)!}{(2\sigma-2)^{2\nu+3}} +$$

$$+\frac{7}{e\log 2}\log T\frac{(2\nu+3)!}{(2\sigma-2)^{2\nu+2}} < T\frac{(2\nu+2)!}{(2\sigma-1)^{2\nu+2}} + 54\log T\,\frac{(2\nu+3)!}{(2\sigma-2)^{2\nu+2}} =$$

$$= T\frac{(2\nu+3)!}{(2\sigma-1)^{2\nu+2}}\left[1+54\log T\,\frac{1}{T}\left(1+\frac{1}{2\sigma-2}\right)^{2\nu+2}\right] <$$

$$< 109\,T\log T\frac{(2\nu+3)!}{(2\sigma-1)^{2\nu+2}}, \qquad \text{w. z. b. w.}$$

Lemma IV. *Es sei die integrierbare Funktion* $A(t)$ *für* $T/2 \leqq t \leqq T$ *definiert und* β *eine beliebige Zahl zwischen* 0 *und* 1. *Dann gilt für* $T/2 \leqq t \leqq T$

$$(26) \qquad |A(t)| \leqq T^{-\frac{\beta}{2}}\left(\int\limits_{T/2}^{T}|A(t)|^2\,dt\right)^{1/2},$$

mit Ausnahme einer Menge, deren Maß höchstens T^{β} *ist.*

Beweis: Wir bezeichnen mit $\mathfrak{R}$ die Menge derjenigen Punkte, für welche (26) nicht erfüllt ist. Dann ist nach dem Mittelwertsatz

$$\int\limits_{T/2}^{T}|A(t)|^2\,dt \geqq m(\mathfrak{R})\inf_{y\in\mathfrak{R}}|A(y)|^2 \geqq m(\mathfrak{R})\,T^{-\beta}\int\limits_{T/2}^{T}|A(t)|^2\,dt,$$

woraus die Behauptung unmittelbar folgt.

Wir wenden Lemma IV mit $A(t) = \frac{\zeta'}{\zeta}(s)^{(\nu)}$, $s = \sigma_0 + it$, $\beta = \alpha$ und Lemma III an. So erhalten wir, daß die Ungleichung

$$\left|\frac{\zeta'}{\zeta}(s)^{(\nu)}\right|_{s=\sigma_0+it} \leqq T^{-\frac{\alpha}{2}}\left(109\,T\log T\frac{(2\nu+3)!}{(2\sigma_0-1)^{2\nu+2}}\right)^{1/2}$$

und nach (16c) und (21) a fortiori die Ungleichung

$$\left|\frac{\zeta'}{\zeta}(s)^{(\nu)}\right|_{s=\sigma_0+it} \leqq 2^{15}\,T^{\frac{1-\alpha}{2}}\log^5 T\,\frac{\nu!}{\left(\sigma_0-\frac{1}{2}\right)^{\nu+1}}$$

für $T > C$ und $T/2 \leqq t \leqq T$ mit Ausnahme einer Menge vom Maß nicht größer als T^{α} gilt. Diese Menge selbst kann wohl von ν abhängen; da aber die Anzahl der von (21) zugelassenen Werte von ν kleiner als $B\log^3 T$ ist, so gilt die folgende Behauptung: für $T > C$ kann man aus dem Intervall $T/2 \leqq t \leqq T$ eine Menge

$\mathfrak{A}^+ = \mathfrak{A}^+(\alpha)$ mit dem Maß $\leqq BT^\alpha \log^3 T$ auslassen, auf deren Komplementärmenge $\mathfrak{B} = \mathfrak{B}(\alpha)$ überall

$$\left| \frac{1}{\nu!} \frac{\zeta'}{\zeta}(s)^{(\nu)} \right|_{s=\sigma_0+it} < 2^{15} T^{\frac{1-\alpha}{2}} \log^5 T \frac{1}{\left(\sigma_0 - \frac{1}{2}\right)^{\nu+1}} \tag{27}$$

für jeden, (21) *erfüllenden Wert* von ν gilt.

Nun betrachten wir auf der Geraden $\sigma = \sigma_0$ die Intervalle $l_\mu \equiv \left[\sigma_0 + \left(\mu + \frac{1}{4}\right) i,\ \sigma_0 + \left(\mu + \frac{3}{4}\right) i\right]$; hier läuft der Index μ die Werte $\left[\frac{T}{2}\right] + 1, \left[\frac{T}{2}\right] + 2, \ldots, [T] - 1$ durch. Nach (27) kann die Menge $\mathfrak{A}^+$ höchstens $2BT^\alpha \log^3 T$ von den Intervallen l_μ ganz bedecken; wir bezeichnen diese mit l''_μ, lassen sie außer Acht und betrachten nur die übrigen Intervalle, welche wir mit l'_μ bezeichnen. Die Anzahl dieser Intervalle l'_μ ist nach dem obigen für $T > C$ größer als $T/2 - 2^7 T^\alpha \log^3 T$ und *jede* l'_μ enthält mindestens eine $s = s_0$, für welche die Ungleichung (27) erfüllt ist. *Im ganzen* § II *bezeichnet* $s_0 = \sigma_0 + it_0$ *durchwegs einen solchen Punkt.*

Aus (11) gewinnen wir nach ν-maliger Differentiation für $\sigma > 1$

$$\frac{(-1)^{\nu+1}}{\nu!} \frac{\zeta'}{\zeta}(s)^{(\nu)} = \frac{1}{(s-1)^{\nu+1}} - \sum_{n=1}^{\infty} \frac{1}{(s+2n)^{\nu+1}} - \sum_{\varrho} \frac{1}{(s-\varrho)^{\nu+1}}.$$

Aus dieser Gleichung und aus (27) erhalten wir für $T > C$

$$\left| \frac{1}{(s_0-1)^{\nu+1}} - \sum_{n=1}^{\infty} \frac{1}{(s+2n)^{\nu+1}} - \sum_{\varrho} \frac{1}{(s_0-\varrho)^{\nu+1}} \right| < 2^{15} \frac{T^{\frac{1-\alpha}{2}} \log^5 T}{\left(\sigma_0 - \frac{1}{2}\right)^{\nu+1}},$$

wo T, α, ν nur durch (18) und (21) beschränkt werden. Wenn ϱ' diejenige ζ-Wurzel bedeutet, welche zu s_0 am nächsten liegt, so gilt

$$\frac{1}{|s_0-\varrho'|^{\nu+1}} \left| \sum_{\varrho} \left(\frac{s_0-\varrho'}{s_0-\varrho}\right)^{\nu+1} + \sum_{n=1}^{\infty} \left(\frac{s_0-\varrho'}{s_0+2n}\right)^{\nu+1} - \left(\frac{s_0-\varrho'}{s_0-1}\right)^{\nu+1} \right| < \tag{28}$$

$$< 2^{15} \frac{T^{\frac{1-\alpha}{2}} \log^5 T}{\left(\sigma_0 - \frac{1}{2}\right)^{\nu+1}}.$$

Zuerst schätzen wir in (28) das letzte Glied ab. Für $T > C$ $(> 2A + 2)$ gilt nach (12)

$$|s_0 - \varrho'|^2 \leqq \sigma_0^2 + \frac{1}{4},$$

folglich nach (20) und (21) und $t_0 \geqq \frac{T}{2}$

$$(29)\quad \left|\frac{s_0-\varrho'}{s_0-1}\right|^{\nu+1} < \left(\frac{\sigma_0^2+\frac{1}{4}}{t_0^2}\right)^{\frac{\nu+1}{2}} < \exp\left(-\frac{\nu}{3}\log T\right) < \exp(-3\log^2 T)$$

für $T>C$. Die Abschätzung der zweiten Summe in (28) gelingt auch sehr einfach; für $T>C$ ist

$$(30a)\quad \begin{aligned} J_6 &\equiv \left|\sum_{n=1}^{\infty}\left(\frac{s_0-\varrho'}{s_0+2n}\right)^{\nu+1}\right| = \left|\frac{s_0-\varrho'}{s_0+2}\right|^{\nu+1}\left|\sum_{n=1}^{\infty}\left(\frac{s_0+2}{s_0+2n}\right)^{\nu+1}\right| < \\ &< \left|\frac{s_0-\varrho'}{s_0-1}\right|^{\nu+1}\sum_{n=1}^{\infty}\left|\frac{s_0+2}{s_0+2n}\right|^{\nu+1} < \exp(-3\log^2 T)\sum_{n=1}^{\infty}\left|\frac{s_0+2}{s_0+2n}\right|^{\nu+1} \end{aligned}$$

nach (29). Ferner gilt für $T>C$, $t_0 \geqq T/2 > 3\sigma_0$, also, wenn $n \geqq l t_0$,

$$4\sigma_0+4 \leqq \sigma_0^2 < \sigma_0^2+t_0^2,$$

$$(\sigma_0+2)^2+t_0^2 < 2(\sigma_0^2+t_0^2) < \frac{20}{9}\,t_0^2 \leqq \frac{3n^2}{l^2} < \frac{(\sigma_0+2n)^2+t_0^2}{l^2}.$$

Daher ist

$$(30b)\quad \begin{aligned} &\sum_{n=1}^{\infty}\left|\frac{s_0+2}{s_0+2n}\right|^{\nu+1} < \\ &< \sum_{1\leqq n\leqq 2t_0} 1 + \frac{1}{2^{\nu+1}}\sum_{2t_0<n\leqq 3t_0} 1 + \frac{1}{3^{\nu+1}}\sum_{3t_0<n\leqq 4t_0} 1 + \ldots < \\ &< 2t_0\left(1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots\right) < 4T. \end{aligned}$$

Aus (28), (29), (30a), (30b) erhalten wir für $T>C$

$$(31)\quad \frac{1}{|s_0-\varrho'|^{\nu+1}}\left(\left|\sum_{\varrho}\left(\frac{s_0-\varrho'}{s_0-\varrho}\right)^{\nu+1}\right| - \exp(-2\log^2 T)\right) < 2^{16}\frac{T^{\frac{1-\alpha}{2}}\log^5 T}{\left(\sigma_0-\frac{1}{2}\right)^{\nu+1}}.$$

Die restliche Summe in (31) zerlegen wir in fünf Teile:

$$(32)\quad \begin{aligned} &J_7 = \sum_{\substack{\varrho \\ t_\varrho \leqq -A}}\left(\frac{s_0-\varrho'}{s_0-\varrho}\right)^{\nu+1}, \quad J_8 = \sum_{-A<t_\varrho\leqq A}\left(\frac{s_0-\varrho'}{s_0-\varrho}\right)^{\nu+1}, \\ &J_9 = \sum_{A<t_\varrho\leqq t_0-2\pi\sqrt{\sigma_0}}\left(\frac{s_0-\varrho'}{s_0-\varrho}\right)^{\nu+1}, \quad J_{10} = \sum_{\substack{\varrho \\ |t_\varrho-t_0|<2\pi\sqrt{\sigma_0}}}\left(\frac{s_0-\varrho'}{s_0-\varrho}\right)^{\nu+1}, \\ &J_{11} = \sum_{\substack{\varrho \\ t_\varrho\geqq t_0+2\pi\sqrt{\sigma_0}}}\left(\frac{s_0-\varrho'}{s_0-\varrho}\right)^{\nu+1} \end{aligned}$$

Diese Fünfteilung hat für $T>C$ gewiß einen Sinn. Für J_8 haben wir gleich wegen (13), (20) und (21)

$$(33a)\qquad |J_8| < \sum_{\substack{\varrho\\ |t_\varrho|\le A}} \left(\frac{\sigma_0^2+\frac{1}{4}}{(\sigma_0-1)^2+(t_0-A)^2}\right)^{\frac{\nu+1}{2}} <$$

$$< \frac{A^2}{2}\left(\frac{2^{18}\log^4 T}{T^2}\right)^{\frac{\nu+1}{2}} < \exp(-2\log^2 T)$$

für $T>C$. Für J_7 kommen wir mit der rohen Abschätzung aus (15) aus; für $T>C$ gilt nämlich

$$|J_7| < \sum_{m=-\infty}^{-A} \sum_{\substack{\varrho\\ -m<t_\varrho\le-(m-1)}} \left|\frac{s_0-\varrho'}{s_0-\varrho}\right|^{\nu+1} < \sum_{m=A}^{\infty} m\left(\frac{\sigma_0^2+\frac{1}{4}}{(\sigma_0-1)^2+(t_0+m)^2}\right)^{\frac{\nu+1}{2}} =$$

$$(33b)\qquad = \left(\frac{\sigma_0^2+\frac{1}{4}}{\sigma_0^2}\right)^{\frac{\nu+1}{2}} \sum_{m=A}^{\infty} m\left(\frac{\sigma_0^2}{(\sigma_0-1)^2+(t_0+m)^2}\right)^{\frac{\nu+1}{2}} <$$

$$< 2\sum_{m=A}^{\infty} m\left(\frac{\sigma_0}{t_0+m}\right)^{\nu+1} < 2\sum_{m=A}^{[t_0^2]+1} m\left(\frac{\sigma_0}{t_0}\right)^{\nu+1} + 2\sum_{m=[t_0^2]+2}^{\infty} m\left(\frac{\sigma_0}{m}\right)^{\nu+1} <$$

$$< 2\left(\frac{\sigma_0}{t_0}\right)^{\nu+1} t_0^5 + 2\,\frac{\sigma_0^{\nu+1}}{t_0^{2\nu-2}} < \exp(-2\log^2 T).$$

Für J_9 haben wir für $T>C$

$$|J_9| < \sum_{m=A}^{[t_0-2\pi\sqrt{\sigma_0}]} m\left(\frac{\sigma_0^2+\frac{1}{4}}{(\sigma_0-1)^2+(m-t_0)^2}\right)^{\frac{\nu+1}{2}} <$$

$$(33c)\qquad < T^2\left(\frac{\sigma_0^2+\frac{1}{4}}{\sigma_0^2+(4\pi^2-2)\sigma_0+1}\right)^{\frac{\nu+1}{2}} < T^2\left(1-\frac{4\pi^2-3}{\sigma_0}\right)^{\frac{\nu+1}{2}} <$$

$$< T^2\exp\left(-\frac{4\pi^2-3}{2}\,\frac{\nu+1}{\sigma_0}\right)$$

und für J_{11}

$$|J_{11}| < \sum_{m=[t_0+2\pi\sqrt{\sigma_0}]}^{\infty} m\left(\frac{\sigma_0^2+\frac{1}{4}}{(\sigma_0-1)^2+(m-t_0)^2}\right)^{\frac{\nu+1}{2}} =$$

$$= \sum_{m=[t_0+2\pi\sqrt{\sigma_0}]}^{[t_0^2]+1} + \sum_{m=[t_0^2]+2}^{\infty} <$$

$$(33d)\qquad < T^4\left(\frac{\sigma_0^2+\frac{1}{4}}{\sigma_0^2+(4\pi^2-2)\sigma_0+1}\right)^{\frac{\nu+1}{2}} + \sum_{m=[t_0^2]+2}^{\infty} m\left(\frac{\sigma_0^2+1}{(m-t_0)^2}\right)^{\frac{\nu+1}{2}} <$$

$$< T^4\left(1-\frac{4\pi^2-3}{\sigma_0}\right)^{\frac{\nu+1}{2}}+\sum_{m=[t_0^3]+2}^{\infty}\left(\frac{2}{m}\right)^{\frac{\nu-1}{2}}<$$
$$< T^4\exp\left(-\frac{4\pi^2-3}{2}\,\frac{\nu+1}{\sigma_0}\right)+\exp(-2\log^2 T).$$

Aus (33a), (33b), (33c), (33d) (32) und (31) erhalten wir für $T>C$

$$(34)\quad \frac{1}{|s_0-\varrho'|^{\nu+1}}\left(\left|\sum_{\substack{\varrho\\ |t_\varrho-t_0|\leqq 2\pi\sqrt{\sigma_0}}}\left(\frac{s_0-\varrho'}{s_0-\varrho}\right)^{\nu+1}\right|-\right.$$
$$\left.-T^4\exp\left(-\frac{4\pi^2-4}{2}\,\frac{\nu+1}{\sigma_0}\right)-\exp(-\log^2 T)\right)<2^{15}\frac{T^{\frac{1-\alpha}{2}}\log^5 T}{\left(\sigma_0-\frac{1}{2}\right)^{\nu+1}}.$$

Zur Abschätzung von J_{10} benötigen wir außer der Hypothese **b** noch das einfache

L e m m a V. *Für die Anzahl N_1 der nichttrivialen Wurzeln im $|t-t_0|\leqq 2\pi\sqrt{\sigma_0}$ gilt für $T>C$*

$$|N_1-2\sqrt{B}\log^2 T|<10\sqrt{B}\log T.$$

B e w e i s: Nach (14) ist für $T>C$

$$\left|N_1-\frac{t_0+2\pi\sqrt{\sigma_0}}{2\pi}\log\frac{t_0+2\pi\sqrt{\sigma_0}}{2\pi}+\right.$$
$$\left.+\frac{t_0-2\pi\sqrt{\sigma_0}}{2\pi}\log\frac{t_0-2\pi\sqrt{\sigma_0}}{2\pi}-2\sqrt{\sigma_0}\right|<\log T,$$

$$\left|N_1-\sqrt{\sigma_0}\log\frac{t_0^2-4\pi^2\sigma_0}{4\pi^2}\right|<$$
$$<(1+2\sqrt{B})\log T+\frac{t_0}{2\pi}\log\frac{1+\frac{2\pi\sqrt{\sigma_0}}{t_0}}{1-\frac{2\pi\sqrt{\sigma_0}}{t_0}}<(1+4\sqrt{B})\log T,$$

$$|N_1-2\sqrt{\sigma_0}\log t_0|<(1+4\sqrt{B})\log T+\sqrt{\sigma_0}\log(4\pi^2)+$$
$$+\sqrt{\sigma_0}\cdot\log\left(\frac{1}{1-\frac{4\pi^2\sigma_0}{t_0^2}}\right)<(1+8\sqrt{B})\log T,$$

$$|N_1-2\sqrt{\sigma_0}\log T|<(1+8\sqrt{B})\log T+\sqrt{\sigma_0}<(1+9\sqrt{B})\log T,\ \text{w.z.b.w.}$$

Wir bemerken zunächst, daß die Glieder $\left(\frac{s_0-\varrho'}{s_0-\varrho}\right)^{\nu+1}$ von J_{10} dem absoluten Werte nach sämtlich $\leqq 1$ sind und Gleichheitzeichen gilt sicher für $\varrho=\varrho'$; ferner, daß die Mengen $\mathfrak{A}^+$, $\mathfrak{B}$, die Intervalle I'_μ, s_0, ϱ', also auch die Zahlen $\frac{s_0-\varrho'}{s_0-\varrho}$ *von der Wahl von ν*

unabhängig sind. Dann können wir aber unsere Hypothese **b** mit $n = N_1$, $z_\nu = \frac{s_0-\varrho'}{s_0-\varrho}$ anwenden, wo N_1 den im Lemma V gegebenen Sinn hat; dieses ergibt für $T > C$ die Existenz einer ganzen Zahl ν_0, für welche

$$N_1^{3/2}(1-N_1^{-0{\cdot}42}) \leqq \nu_0 \leqq N_1^{3/2} \tag{35a}$$

und

$$|J_{10}| = \Bigg| \sum_{\substack{\varrho \\ |t_\varrho - t_0| \leqq 2\pi\sqrt{\sigma_0}}} \left(\frac{s_0-\varrho'}{s_0-\varrho}\right)^{\nu_0} \Bigg| > \exp(-N_1^{0{\cdot}09}). \tag{35b}$$

Wir geben nun dem, bisher nur der Bedingung (21) untergeworfenen ν den Wert $\nu = \nu_0 - 1$; wir zeigen, daß (21) bei dieser Wahl befriedigt wird. Aus (35a) und Lemma V folgt für $T > C$

$$\nu < \nu_0 \leqq N_1^{3/2} < (2\sqrt{B}\log^2 T)^{3/2}\left(1+\frac{5}{\log T}\right)^{3/2} < \tag{36a}$$
$$< 2^{3/2}B^{3/4}\log^3 T\left(1+\frac{8}{\log T}\right) < B\log^3 T$$

nach der Definition von B (siehe (20)). Es folgt ferner aus Lemma V und (35a) ganz grob für $T > C$

$$\nu + 1 = \nu_0 > \frac{1}{2}N_1^{3/2} > \frac{1}{2}(\sqrt{B}\log^2 T)^{3/2} > 2B\log^2 T + 1. \tag{36b}$$

Mit (36a) und (36b) ist also (21) gerechtfertig. Da für $T > C$ offenbar

$$\frac{\nu_0}{\sigma_0} > \frac{2}{3\sigma_0}N_1^{3/2} > \frac{2}{3}\,\frac{2^{3/2}B^{3/4}\log^3 T\left(1-\frac{5}{\log T}\right)^{3/2}}{B\log^2 T} > \frac{1}{2}\log T$$

und

$$T^4\exp\left(-\frac{4\pi^2-4}{2}\,\frac{\nu+1}{\sigma_0}\right) = T^4\exp\left(-\frac{4\pi^2-4}{2}\,\frac{\nu_0}{\sigma_0}\right) <$$
$$< T^{4-(\pi^2-1)} < T^{-4}$$

gilt, erhalten wir aus diesen, (35b), (34) und Lemma V für $T > C$

$$|s_0-\varrho'|^{-\nu_0}\exp(-2\log^{0{\cdot}18} T) <$$
$$< \frac{1}{|s_0-\varrho'|^{\nu+1}}\{\exp(-N_1^{0{\cdot}09}) - 2T^{-4}\} < 2^{15}\frac{T^{\frac{1-\alpha}{2}}\log^5 T}{\left(\sigma_0-\frac{1}{2}\right)^{\nu_0}},$$

d. h. für $T > C$

$$|s_0 - \varrho'| >$$
$$> \left(\sigma_0 - \frac{1}{2}\right)\exp\left(-\frac{1-\alpha}{2\nu_0}\log T - \frac{5\log\log T}{\nu_0} - \frac{15\log 2}{\nu_0} - \frac{2\log^{0\cdot18} T}{\nu_0}\right) >$$
$$> \left(\sigma_0 - \frac{1}{2}\right)\left[1 - \left(\frac{1-\alpha}{2}\log T + 5\log\log T + 15\log 2 + 2\log^{0\cdot18} T\right)\frac{1}{\nu_0}\right] >$$
$$(37) \qquad > \left(\sigma_0 - \frac{1}{2}\right)\left[1 - \left(\frac{1-\alpha}{2}\log T + 3\log^{0\cdot18} T\right)\frac{1}{\nu_0}\right] >$$
$$> \sigma_0 - \frac{1}{2} - \left(\frac{1-\alpha}{2}\log T + 3\log^{0\cdot18} T\right)\frac{\sigma_0}{\nu_0}.$$

Mit Rücksicht auf (35a), (20) und Lemma V ist für $T > C$

$$\frac{\sigma_0}{\nu_0} < \frac{B\log^2 T}{N_1^{3/2}[1 - N_1^{-0\cdot42}]} < \frac{B\log^2 T}{(2\sqrt{B})^{3/2}\log^3 T\left(1 - \frac{5}{\log T}\right)^{3/2}(1 - N_1^{-0\cdot42})} <$$
$$< \frac{(1 + \log^{-0\cdot84} T)^{1/4}}{\log T}\left(1 + \frac{8}{\log T}\right)\left(1 + \frac{2}{N_1^{0\cdot42}}\right) < \frac{1}{\log T}\left(1 + \frac{1}{\log^{0\cdot84} T}\right),$$

also folgt aus (37)

$$(38) \qquad |s_0 - \varrho'| > \sigma_0 - \left(1 - \frac{\alpha}{2}\right) - \frac{4}{\log^{0\cdot82} T}.$$

Nun betrachten wir von den Intervallen I'_μ ein beliebiges I'_j und wir bezeichnen durch $\varrho^+ = \sigma^+ + it^+$ diejenige nichttriviale Wurzel aus $j \leqq t_\varrho \leqq j+1$, welche den größten reellen Teil besitzt. Dann folgt aus (38) für $T > C$

$$\sigma_0^2 - 2\sigma^+\sigma_0 + 2 \geqq (\sigma_0 - \sigma^+)^2 + (t_0 - t^+)^2 = |s_0 - \varrho^+|^2 \geqq |s_0 - \varrho'|^2 >$$
$$> \sigma_0^2 - 2\left(1 - \frac{\alpha}{2}\right)\sigma_0 - \frac{8\sigma_0}{\log^{0\cdot82} T},$$

d. h.

$$(39) \qquad \sigma^+ \leqq 1 - \frac{\alpha}{2} + \frac{5}{\log^{0\cdot82} T}.$$

Dies bedeutet aber, daß im Intervall $T/2 \leqq t \leqq T$ nur die Intervalle I''_μ nichttriviale, in der Halbebene $\sigma \geqq 1 - \frac{\alpha}{2} + \frac{5}{\log^{0\cdot82} T}$ liegende Wurzeln enthalten können, wenn nur $T > C$. Daraus folgt nach (15) und $m(\mathfrak{A}^+) \leqq BT^\alpha \log^3 T$ für $T > C$

$$N^+\left(1 - \frac{\alpha}{2} + \frac{5}{\log^{0\cdot82} T}, T\right) < 2^7 T^\alpha \log^4 T.$$

Wenn wir α durch $\alpha^+ + \frac{10}{\log^{0\cdot 82} T}$ ersetzen, gelangen wir zu

$$(40) \qquad N^+\left(1-\frac{\alpha^+}{2}, T\right) < 2^7 T^{[\alpha^+ + 10\exp(-0\cdot 82 \log\log T)]} \log^4 T <$$

$$< T^{\alpha^+} \exp(11 \log^{0\cdot 18} T),$$

wenn $T > C$ und $1 - \frac{10}{\log^{0\cdot 82} T} \geqq \alpha^+ \geqq \frac{2}{\log^{0\cdot 81} T}$; nach (14) gilt diese Abschätzung auch für $1 \geqq \alpha^+ \geqq 1 - \frac{10}{\log^{0\cdot 82} T}$. Wir bezeichnen α^+ wieder durch α und wenden (40) nacheinander mit $\frac{T}{2}, \frac{T}{2^2}, \ldots, \frac{T}{2^\lambda}$ anstatt T an, wo

$$\lambda = \left[\left(1-\frac{\alpha}{4}\right)\frac{\log T}{\log 2}\right] + 1.$$

Dann ist offenbar

$$\lambda < 2 \log T, \ \frac{T}{2^\lambda} \leqq T^{\frac{\alpha}{4}} \leqq \frac{T}{2^{\lambda-1}};$$

für $\alpha \geqq \frac{2}{\log^{0\cdot 81} T}$ ist

$$\frac{T}{2^\lambda} > \frac{1}{2} T^{\frac{\alpha}{4}} > \frac{1}{2} \exp\left(\frac{1}{2} \log^{0\cdot 19} T\right) > C$$

wenn $T > C$. Nach Summation erhalten wir für $T > C$

$$N\left(1-\frac{\alpha}{2}, T\right) - N\left(1-\frac{\alpha}{2}, T^{\frac{\alpha}{4}}\right) <$$

$$< T^\alpha \exp(11 \log^{0\cdot 18} T)\left[\frac{1}{1^\alpha} + \frac{1}{2^\alpha} + \ldots + \frac{1}{\lambda^\alpha}\right] > T^\alpha \exp(12 \log^{0\cdot 18} T)$$

und, da nach (15) offenbar $N\left(1-\frac{\alpha}{2}, T^{\frac{\alpha}{4}}\right) < 2 T^{\frac{\alpha}{4}} \log T$ gilt, gewinnen wir endlich für $T > C$, $1 \geqq \alpha \geqq 2 \log^{-0\cdot 81} T$ die behauptete Ungleichung (17).

Für den Beweis der asymptotischen Formel (9) skizzieren wir vollständigkeitshalber die Beweisführung von HOHEISEL—INGHAM. Zur Abkürzung sei gesetzt:

$$\sqrt{x} \exp(\log^{0\cdot 996} x) = h, \quad \sum_{\substack{p,\, m \\ p^m \leqq x}} \log p = \psi(x), \quad T = \sqrt{x} \exp(-\log^{0\cdot 995} x).$$

Aus der Riemann—Mangoldt—Landauschen Formel erhalten die genannten Verfasser leicht (siehe [8]))

(41) $$\left|\frac{\psi(x+h)-\psi(x)}{h}-1\right|<$$

$$< C\left(\frac{x}{Th}\log^2 x+\frac{T\log T}{x}+2\int_0^{1-2\log^{-0.81}T} N(a,T)x^{a-1}\log x\,da\right),$$

wo C von da an von x unabhängige Konstanten bedeutet. Die ersten zwei Glieder haben den Grenzwert 0; für das dritte gilt für $T>C$ nach (5)

$$\int_0^1 N(a,T)x^{a-1}\log x\,da<$$

$$<\frac{10T\log T\log x}{\sqrt{x}}+\int_{1/2}^{1-2\log^{-0.81}T} N(a,T)x^{a-1}\log x\,da<$$

$$<o(1)+\log x\int_{-\infty}^{1-2\log^{-0.81}T}\left(\frac{T^2}{x}\right)^{1-a}\exp(13\log^{0.18}T)\,da<$$

$$<o(1)+\exp(14\log^{0.18}T)\exp[2\log^{-0.81}T(-2\log^{0.995}x)]<$$

$$<o(1)+\exp(14\log^{0.18}T)\exp(-\log^{0.184}T)=o(1),$$

d. h.

$$\frac{\psi(x+h)-\psi(x)}{h}=1+o(1),\qquad \text{w. z. b. w.}$$

§ III.

Dieser Abschnitt enthält einige Bemerkungen über unsere Hypothese **a**. Wir beweisen den folgenden

S a t z. *Für* $m\geqq 2n$, $|z_\nu|\geqq 1-\frac{n}{m}$ $(\nu=1,2,\ldots,n)$ *gilt*

$$R=\max_{m-n+1\leqq\nu\leqq m}|z_1^\nu+z_2^\nu+\ldots+z_n^\nu|>\left(\frac{n}{200m}\right)^n.$$

B e w e i s: Es sei

(42a) $$f(z)=\prod_{\nu=1}^{n}\left(1-\frac{z}{z_\nu}\right)=\sum_{\nu=0}^{n}b_\nu z^\nu,\qquad b_0=1$$

(42b) $$\frac{1}{f(z)}=\sum_{\nu=0}^{\infty}c_\nu z^\nu,\qquad c_0=1.$$

Offenbar gilt

$$(42c)\qquad c_u = \sum_{\substack{i_1+i_2+\dots+i_n=\mu \\ i_1\geqq 0,\, i_2\geqq 0,\dots,\, i_n\geqq 0}} \frac{1}{z_1^{i_1}, z_2^{i_2}, \dots, z_n^{i_n}},$$

also

$$(43a)\qquad |b_\nu| \leqq \binom{n}{\nu} \frac{1}{\left(1-\frac{n}{m}\right)^\nu} \qquad (\nu = 0, 1, 2, \dots, n)$$

und

$$(43b)\qquad |c_\mu| \leqq \frac{1}{\left(1-\frac{n}{m}\right)^\mu} \sum_{\substack{j_1+j_2+\dots+j_n=n+\mu \\ j_1>0,\dots,\, j_n>0}} 1 = \frac{1}{\left(1-\frac{n}{m}\right)^\mu} \binom{\mu+n-1}{n-1}.$$

Es sei

$$(44a)\qquad \sum_{\nu=0}^{m-n} c_\nu z^\nu = s_{m-n}(z)$$

und

$$(44b)\qquad 1 - f(z) s_{m-n}(z) = F(z).$$

Offenbar ist $F(z)$ ein Polynom m-ten Grades; es ist leicht zu sehen, daß die Koeffizienten von $z^0, z^1, z^2, \dots, z^{m-n}$ in $F(z)$ verschwinden. Es gilt

$$(45a)\qquad F(z) = \sum_{\nu=m-n+1}^{m} d_\nu z^\nu$$

und nach (44b)

$$(45b)\qquad F(z_1) = F(z_2) = \dots = F(z_n) = 1.$$

Wir ersetzen in (45a) z der Reihe nach durch $z_1, z_2, \dots, z_n$ und addieren; so gelangen wir wegen (45b) mit der Abkürzung $z_1^\nu + z_2^\nu + \dots + z_n^\nu = s_\nu$ zu

$$(46)\qquad n = \sum_{\nu=m-n+1}^{m} d_\nu s_\nu = \left|\sum_{\nu=m-n+1}^{m} d_\nu s_\nu\right| \leqq R \sum_{\nu=m-n+1}^{m} |d_\nu|.$$

Da für $k = (m-n+1), (m-n+2), \dots, m$ nach (43a) und (43b) wegen $m > 2n$

$$|d_k| = |c_k b_0 + c_{k-1} b_1 + \dots + c_{k-n} b_n| \leqq$$

$$(47a)\qquad \leqq \frac{1}{\left(1-\frac{n}{m}\right)^k}\left[\binom{k+n-1}{n-1}\binom{n}{0} + \binom{k+n-2}{n-1}\binom{n}{1} + \dots + \binom{k-1}{n-1}\binom{n}{n}\right] \leqq \frac{\binom{k+n-1}{n-1} 2^n}{\left(1-\frac{n}{m}\right)^m} \leqq \frac{2^n}{\left(1-\frac{n}{m}\right)^m}\binom{m+n-1}{n-1}$$

und wegen der elementaren Ungleichung $(n-1)^{n-1} < e^{n-2}(n-1)!$

$$\binom{m+n-1}{n-1} < \frac{(m+n)^{n-1}e^{n-2}}{(n-1)^{n-1}} < \left(e\frac{m+n}{n-1}\right)^n < \left(3e\frac{m}{n}\right)^n,$$

gilt, so folgt

$$\text{(47b)} \qquad |d_k| \leqq \frac{\left(6e\frac{m}{n}\right)^n}{\left(1-\frac{n}{m}\right)^m} \qquad (k=m-n+1,\ldots,m).$$

Nach (47b) und (46) ist

$$R \geqq \frac{\left(1-\frac{n}{m}\right)^m}{\left(6e\frac{m}{n}\right)^n} > \frac{e^{-2n}}{\left(6e\frac{m}{n}\right)^n} = \left(\frac{n}{6e^3m}\right)^n > \left(\frac{n}{200m}\right)^n, \quad \text{w. z. b. w.}$$

Endlich zeigen wir an einem Gegenbeispiel, daß es zu jedem M ein Wertsystem $(z_1,\ldots,z_n)$ sogar mit $z_1=1=|z_2|=\ldots=|z_n|$ gibt, für welche

$$\max_{M\left(1-\frac{1}{\omega}\right)\leqq\nu\leqq M} |z_1^\nu+z_2^\nu+\ldots+z_n^\nu| < \exp\left(-\frac{5}{6}\log n \log\omega\right).$$

Das Gegenbeispiel habe ich aus einer Bemerkung des Herrn P. Stein gewonnen, die ich aus einem Littlewoodschen Aufsatze[15]) entnehme; dieser Aufsatz behandelt übrigens ein Problem, welches mit unseren Hypothesen verwandt ist. Es sei $n=4^l$, die ersten $\binom{2l}{l}$ von den z_ν seien gleich 1, die folgenden $\binom{2l}{l-1}$ aus denselben seien gleich $\exp\left(\frac{\pi i}{M}\right)$, die folgenden $\binom{2l}{l+1}$ seien gleich $\exp\left(-\frac{\pi i}{M}\right)$ und allgemein

$$z_{\binom{2l}{l}+\binom{2l}{l-1}+\binom{2l}{l+1}+\ldots+\binom{2l}{l-k+1}+\binom{2l}{l+k-1}+1} = \ldots =$$
$$= z_{\binom{2l}{l}+\binom{2l}{l-1}+\binom{2l}{l+1}+\ldots+\binom{2l}{l-k+1}+\binom{2l}{l+k-1}+\binom{2l}{l-k}} = \exp\left(\frac{k\pi i}{M}\right)$$

$$z_{\binom{2l}{l}+\binom{2l}{l-1}+\binom{2l}{l+1}+\ldots+\binom{2l}{l-k}+1} = \ldots =$$
$$= z_{\binom{2l}{l}+\binom{2l}{l-1}+\binom{2l}{l+1}+\ldots+\binom{2l}{l-k+1}+\binom{2l}{l+k-1}+\binom{2l}{l-k}+\binom{2l}{l+k}} = \exp\left(-\frac{k\pi i}{M}\right)$$
$$(k=2,3,\ldots,l).$$

---

15) J. E. Littlewood, Mathematical Notes (12): An inequality for a sum of cosines, *Journal of London Math. Society*, **12** (1937), p. 117—222.

Dann ist

$$z_1^x + z_2^x + \ldots + z_n^x = \sum_{\mu=0}^{2l} \binom{2l}{\mu} \left(e^{(l-\mu)\frac{\pi i}{M}}\right)^x = \sum_{\mu=0}^{2l} \binom{2l}{\mu} \left(e^{\frac{\pi i x}{M}}\right)^{l-\mu} =$$

$$= \left(e^{\frac{\pi i x}{2M}} + e^{-\frac{\pi i x}{2M}}\right)^{2l} = \left(2\cos\frac{\pi x}{2M}\right)^{2l},$$

also

$$\max_{M\left(1-\frac{1}{\omega}\right) \leqq \nu \leqq M} |z_1^\nu + \ldots + z_n^\nu| \leqq \left(2\cos\left(\frac{\pi}{2} - \frac{\pi}{2\omega}\right)\right)^{2l} =$$

$$= \left(2\sin\frac{\pi}{2\omega}\right)^{2l} < \left(\frac{\pi}{\omega}\right)^{2l} = \exp\left(-\frac{\log n}{\log 2}\log\frac{\omega}{\pi}\right) <$$

$$< \exp\left(-\frac{5}{6}\log n \log\omega\right),$$

wie behauptet.

(Eingegangen am 27. Dezember 1939; umgearbeitet am 1. Januar 1941.)

Eine Bemerkung zu meiner Arbeit „Über die Summabilität der Fourierschen Reihe"[1]).

Von GÉZA GRÜNWALD in Budapest.

1. Es sei $F(\xi, \eta)$ eine L-integrierbare und nach 2π periodische Funktion der reellen Variablen ξ und η, und es bezeichne s_{mn} die Partialsumme vom Index mn der Fourierschen Doppelreihe dieser Funktion an der Stelle (ξ, η). Bekanntlich gibt es L-integrierbare Funktionen, deren arithmetische Mittelwerte

$$\sigma_{mn} = \frac{1}{(m+1)(n+1)} \sum_{\nu=0}^{m} \sum_{\mu=0}^{n} s_{\nu\mu} \tag{1}$$

in sämtlichen Punkten des Quadrats $-\pi \leqq \xi \leqq \pi$, $-\pi \leqq \eta \leqq \pi$ divergieren[2]). Wir können die Konvergenz der arithmetischen Mitteln dadurch sichern, daß wir bei der Mittelbildung einige Partialsummen von s_{mn} ausschließen. So bewies MARCINKIEWICZ und ZYGMUND, daß die Folge (1) fast überall konvergiert, wenn $m/n \leqq \lambda$, $n/m \leqq \lambda$ ist, wo λ eine feste Zahl ist[3]). Noch einfacher ist diejenige Summationsmethode, welche von L. FEJÉR stammt, die von der zweifach unendlichen Folge der Partialsummen s_{mn} die einfach unendliche Folge $s_{00}, s_{11}, \ldots, s_{nn}, \ldots$ auswählt und deren Mitteln untersucht[4]). Wir haben bewiesen, daß diese Mittelwerte für L-integrierbare $F(\xi, \eta)$ fast überall konvergieren[1]). Der Beweis dieses

[1]) *Diese Acta*, **10** (1941), S. 55–63. Wir zitieren im Folgenden diese Arbeit als I.

[2]) S. SAKS, Remark on the differentiability of the Lebesgue indefinite integral, *Fundamenta Math.*, **22** (1934), p, 257–261.

[3]) J. MARCINKIEWICZ and A. ZYGMUND, On the summability of double Fourier series, *Fundamenta Math.*, **32** (1939), p. 122–132.

[4]) L. FEJÉR, Zur Summabilitätstheorie der Fourierschen und Laplaceschen Reihe, *Proceedings Cambridge Phil. Society*, **34** (1938), p. 503–509.

Satzes ist auf den Fall eines anderen einfachen und interessanten Summationsverfahrens übertragbar. Betrachten wir nämlich diejenigen Partialsummen, in denen die Summe der Indizes gleich n ist, also die Partialsummen

$$s_{0n}, s_{1,n-1}, \ldots, s_{n0}; \tag{2}$$

dann gilt für diese der auf die „Cauchysche" Partialsummen bezügliche Summabilitätssatz:

Es sei $F(\xi, \eta)$ eine L-integrierbare und nach 2π periodische Funktion der reellen Variabeln ξ und η. Bezeichnet dann $\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} A_{mn}$ die Fouriersche Doppelreihe dieser Funktion $F(\xi, \eta)$ an der Stelle (ξ, η), so ist fast überall im Quadrat $-\pi \leqq \xi \leqq \pi$, $-\pi \leqq \eta \leqq \pi$

$$\lim_{n=\infty} \frac{s_{0n} + s_{1,n-1} + \ldots + s_{n0}}{n+1} = F(\xi, \eta), \tag{3}$$

wo

$$s_{k,n-k} = \sum_{\nu=0}^{k} \sum_{\mu=0}^{n-k} A_{\nu\mu}. \tag{4}$$

2. Für den Beweis ist es hinreichend zu zeigen, daß für den hier auftretenden Kern die Abschätzungen (16), (17), (18), (19) von I gelten; alles Weitere läßt sich wörtlich aus I übernehmen.

Es sei

$$\bar{\sigma}_n(\xi, \eta; F) = \frac{s_{0,n-1} + s_{1,n-2} + \ldots + s_{n-1,0}}{n}, \tag{5}$$

dann ist

$$\begin{aligned} &\bar{\sigma}_n(\xi, \eta; F) = \\ &= \frac{1}{4\pi^2 n} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} F(u+\xi, t+\eta) \sum_{k=1}^{n} \frac{\sin(k-1/2)t \sin(n-k+1/2)u}{\sin t/2 \sin u/2} du\, dt = \\ &\qquad = \frac{1}{\pi^2} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} F(u+\xi, t+\eta)\, \bar{k}_n(u, t)\, du\, dt. \end{aligned} \tag{6}$$

Wegen

$$\begin{aligned} &\sum_{k=1}^{n} \sin(k-1/2)\, t \sin(n-(k-1/2))\, u = \\ &= 1/2 \cos n u \sum_{k=1}^{n} (\cos(k-1/2)(t+u) - \cos(k-1/2)(t-u)) + \\ &\qquad + 1/2 \sin n u \sum_{k=1}^{n} (\sin(k-1/2)(t+u) + \sin(k-1/2)(t-u)) = \end{aligned}$$

$$= {}^1/_2 \cos nu \left(\frac{\sin n(t+u)}{2\sin\frac{t+u}{2}} - \frac{\sin n(t-u)}{2\sin\frac{t-u}{2}} \right) +$$

$$+ {}^1/_2 \sin nu \left(\frac{\sin^2 n\frac{t+u}{2}}{\sin\frac{t+u}{2}} - \frac{\sin^2 n\frac{t-u}{2}}{\sin\frac{t-u}{2}} \right) =$$

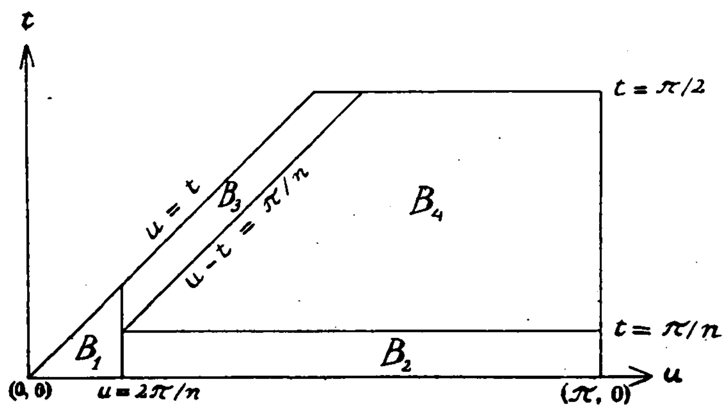
$$= {}^1/_2 \left(\frac{\sin n\frac{t+u}{2}\cos n\frac{t-u}{2}}{\sin\frac{t+u}{2}} - \frac{\sin n\frac{t-u}{2}\cos n\frac{t+u}{2}}{\sin\frac{t-u}{2}} \right)$$

erhalten wir

(7) $\bar{k}_n(u,t) =$

$$= \frac{1}{8n\sin{}^u/_2 \sin{}^t/_2} \left(\frac{\sin n\frac{t+u}{2}\cos n\frac{t-u}{2}}{\sin\frac{t+u}{2}} - \frac{\sin n\frac{t-u}{2}\cos n\frac{t+u}{2}}{\sin\frac{t-u}{2}} \right).$$

Wir werden die Funktion $\bar{k}_n(u, t)$ in den Bereichen B_1, B_2, B_3, B_4 (siehe die Figur) abschätzen.



a) In B_1 ist[5]) wegen (6)

(8) $$|\bar{k}_n(u, t)| \leqq {}^c/_n \sum_{k=1}^{n} (k - {}^1/_2)(n - k + {}^1/_2) \leqq cn^2$$

b) In B_2 ist wegen $u - t \geqq {}^\pi/_n$ und $u - t \geqq {}^u/_2$

$$\left| \frac{\sin n\frac{t+u}{2}}{\sin\frac{t+u}{2}} - \frac{\sin n\frac{t-u}{2}}{\sin\frac{t-u}{2}} \right| \leqq c\,\frac{nt}{u}$$

[5]) Hier und im Folgenden bezeichnen wir einfachheitshalber die verschiedenen absoluten Konstanten mit c, also ohne Indizes.

und

$$\left|\frac{\sin n\frac{t+u}{2}\cos n\frac{t-u}{2}}{\sin\frac{t+u}{2}}-\frac{\sin n\frac{t-u}{2}\cos n\frac{t+u}{2}}{\sin\frac{t-u}{2}}\right|\leqq$$

$$\leqq\left|\frac{\sin n\frac{t+u}{2}}{\sin\frac{t+u}{2}}-\frac{\sin n\frac{t-u}{2}}{\sin\frac{t-u}{2}}\right|+$$

$$+|\sin nt/2\sin nu/2|\left|\frac{\sin n\frac{t+u}{2}}{\sin\frac{t+u}{2}}+\frac{\sin n\frac{t-u}{2}}{\sin\frac{t-u}{2}}\right|\leqq$$

$$\leqq c\frac{nt}{u}+|\sin nu/2\sin nt/2|\left(c\frac{nt}{u}+2\left|\frac{\sin n\frac{t-u}{2}}{\sin\frac{t-u}{2}}\right|\right)\leqq$$

$$\leqq cn\frac{t}{u}+c\frac{|\sin nt/2|}{u}\leqq c\frac{nt}{u},$$

also

$$(9)\qquad |\bar{k}_n(u,t)|\leqq c\,\frac{1}{nut}\,\frac{nt}{u}=c\frac{1}{u^2}.$$

c) In B_3 gilt

$$(10)\qquad |\bar{k}_n(u,t)|\leqq c\,\frac{1}{nut}\,n=c\frac{1}{ut}.$$

d) Endlich ist in B_4

$$(11)\qquad |\bar{k}_n(u,t)|\leqq c\,\frac{1}{nut}\left(\frac{1}{u+t}+\frac{1}{u-t}\right)\leqq c\,\frac{1}{nut(u-t)}.$$

Die Abschätzungen (8), (9), (10), (11) sind mit den Abschätzungen (16), (17), (18), (19) von I identisch, qu. e. d.

(Eingegangen am 9. März 1940)

Über den Fundamentalsatz der Abelschen Gruppen von endlicher Ordnung.

Von L. RÉDEI in Szeged.

Der verblüffend kurze Beweis von KORSELT—FRANZ[1]) für den im Titel genannten Satz erfordert vom Leser wegen der darin vorkommenden künstlichen Doppelminimalannahme anstrengendes Nachdenken, bietet auch wenig Einblick in die wahren Verhältnisse Abelscher Gruppen und keine Vorschrift zur Konstruktion einer Basis. VAN DER WAERDEN[2]) schließt dagegen immer noch kurz genug, auf den Begriff der Faktorgruppen stützend, wobei man den Eindruck hat, daß der passendste Weg gegangen wurde. Will man aber Faktorgruppen vermeiden — ein Standpunkt, der unter Umständen zweckmäßig sein kann — so scheint folgender Beweis unter den zahlreichen anderen von ähnlicher Natur von Vorteil zu sein.

Es sei G eine Abelsche Gruppe von endlicher Ordnung. Es bezeichne (G) die Ordnung von G, (A) die eines Elements A von G, E das Einheitselement von G. Die von E verschiedenen Elemente $A_1, \ldots, A_r$ von G nennen wir eine Basis von G, wenn sie alle von Primzahlpotenzordnung sind, alle Elemente von G in der Form

$$(1) \qquad A_1^{x_1} \ldots A_r^{x_r} \qquad (x_i = 0, 1, \ldots, (A_i)_r - 1\,;\, i = 1, 2, \ldots, r)$$

darstellbar sind und auch

$$(2) \qquad (G) = (A_1) \ldots (A_r)\ ^3)$$

[1]) W. FRANZ, Zur vorstehenden Arbeit von Herrn A. Korselt, *Journal für Math.*, **164** (1931), S. 63.

[2]) B. L. VAN DER WAERDEN, *Moderne Algebra* II (Berlin, 1940), S. 116.

[3]) Offenbar läßt sich (2) mit der Bedingung ersetzen, daß die Produkte (1) verschieden sind, wodurch die übliche Definition der Basis entsteht. Mit der obigen Form der Definition kommen wir ökonomischer aus.

gilt. Der Fundamentalsatz lautet dann: Ist $G \neq E$, so hat G eine Basis.

Dem Beweis schicken wir die aus der Definition unmittelbar folgende Bemerkung voran:

a. Ist $A_1, \ldots, A_r$ Basis einer Gruppe und sind $x_1, \ldots, x_r$ beliebige ganze rationale Zahlen, so ist (1) nur für $(A_i)|x_i$ $(i=1,\ldots,r)$ gleich E.

Es sei G_0 eine Untergruppe von G mit möglichst großem (G_0), die eine Basis $B_1, \ldots, B_s$ hat. Ein solches G_0 existiert sicher, denn G hat ein Element von Primzahlordnung, das dann Basis der von ihm erzeugten Gruppe ist. Man darf $G_0 \neq G$ annehmen, da sonst der Satz richtig ist. Dann gibt es ein Element A in G, das in G_0 nicht enthalten ist. Da jedes Element als ein Produkt von Elementen von Primzahlpotenzordnung darstellbar ist, so läßt sich annehmen, daß die Ordnung von A die Potenz einer Primzahl p ist. Auch läßt sich annehmen, daß A^p in G_0 enthalten ist, da man sonst statt A das letzte Element der Folge $A^p, A^{p^2}, A^{p^3}, \ldots, E$ nehmen könnte, das in G_0 noch nicht enthalten ist. Dann bestehen die Nebengruppen $G_0, AG_0, A^2G_0, \ldots, A^{p-1}G_0$ aus allen verschiedenen Elementen von $\{A, G_0\}$, wobei $\{\ \}$ das Zeichen für die durch die eingeklammerten Elemente erzeugte Gruppe ist. Also gilt

$$(3) \qquad (\{A, G_0\}) = p(B_1) \ldots (B_s).$$

Es ist mit einer geeigneten Reihenfolge von $B_1, \ldots, B_s$

$$(4) \qquad A^p = B_1^{y_1} \ldots B_t^{y_t} \qquad (0 \leqq t \leqq s),$$

wobei die Potenzen rechts von E verschieden sind ($t=0$ bedeutet, daß die rechte Seite gleich E ist). Erhebt man (4) zur (A)-ten Potenz, so folgt nach **a** $(B_i)|(A)y_i$ $(i=1,\ldots,t)$. Wegen $(B_i) \nmid y_i$ ist also $p|(B_i)$, d. h. die Ordnungen von $B_1, \ldots, B_t$ sind Potenzen von p. Es genügt weiter den Fall zu betrachten, in dem kein y_i durch p teilbar ist, denn ist z. B. $p|y_t$, so ist auch $A' = AB_t^{-\frac{y_t}{p}}$ ein in G_0 nicht enthaltenes Element, wofür A'^p in G_0 enthalten ist, uns dann läßt sich A von vornherein durch A' ersetzen, wodurch man in ein paar Schritten zum Ziele kommt.

Ist dann $t=0$, so ist $(A)=p$. Das bedeutet wegen (3), daß $A, B_1, \ldots, B_s$ eine Basis von $\{A, G_0\} = \{A, B_1, \ldots, B_s\}$ ist, das ein Widerspruch ist.

Ist dagegen $t > 0$, so wählt man eine solche Reihenfolge von $B_1, \ldots, B_t$, daß die Ordnung von B_1 möglichst groß ausfällt. Nach (4) und **a** ist $A^{p(B_1)} = E$, $A^{(B_1)} \neq E$, woraus $(A) = p(B_1)$ folgt. Wegen $p \nmid y_1$ gibt es ein ganzes rationales z mit $y_0 z \equiv 1$ (mod (B_1)) und dann ist nach (4) $B_1 = (B_1^{y_0})^z$ in $\{A, B_2, \ldots, B_t\}$ enthalten. Hieraus folgt, daß diese Gruppe gleich $\{A, G_0\}$ und also nach (3) $A, B_2, \ldots, B_t$ eine Basis von ihr ist. Dieser Widerspruch beweist den Satz.

(Eingegangen am 14. August 1941.)

Remarques sur la loi des erreurs.

Par CHARLES JORDAN à Budapest.

§ 1. Certaines questions concernant la loi des erreurs ne sont pas encore complètement élucidées. Il y a des personnes qui pensent que la loi des erreurs peut être démontrée, et qu'en l'appliquant il est possible de découvrir les erreurs dites systématiques, c'est-à-dire dues aux imperfections des instruments de mesure. Or il est évident qu'un fait d'expérience tel que "les erreurs d'observation suivent une certaine loi" ne peut être prouvé par aucune déduction mathématique. L'expérience seule peut en décider, mais là encore il y a des difficultés; en effet, les faits d'expérience, justement à cause des erreurs d'observation inévitables, ne peuvent être conformes d'une manière absolue à aucune loi précise, ils ne peuvent la suivre que d'une manière approchée. Une autre difficulté consiste en ce que pour vérifier la loi il faudrait connaitre la grandeur à mesurer, or cette grandeur est presque toujours inconnue. Dans l'exemple classique de la somme des angles d'un triangle, on mesure les *angles* et on montre que leur *somme* suit la loi; la démonstration n'est pas directe. Pour la compléter il faut remarquer, (on le verra au § 4) que si l'on admet pour les erreurs la loi du hasard, il en résulte que la somme des erreurs la suit aussi. L'expérience fournira tout au plus la conclusion: Dans certaines conditions, les erreurs d'observation suivent d'assez près la loi dite du hasard.

Toutes les observations sont nécessairement affectées d'erreurs. En effet, si on mesure plusieurs fois successivement une même grandeur, les résultats seront différents et les différences seront dues aux erreurs d'observation. En général, on classe les erreurs en deux groupes, dans le premier on range les erreurs dites systématiques, c'est-à-dire produites par des causes constantes, agissant toujours dans le même sens et causées, par exemple, par les imperfections des instruments ; l'observateur doit chercher à éliminer ces causes, car la théorie mathématique des erreurs ne peut rien dire sur ces erreurs.

Dans le second groupe, on classe les erreurs dites fortuites ou accidentelles, c'est-à-dire produites par de petites causes très nombreuses agissant tantôt dans un sens, tantôt dans un autre ; une telle erreur peut être considérée comme la résultante d'un très grand nombre de petites erreurs élémentaires n'ayant chacune que très peu d'influence sur le résultat. Une réflexion simple montre que ces dernières peuvent suivre une loi mathématique. En effet si l'on suppose par exemple que chaque cause élémentaire peut produire avec la même facilité une erreur $+1$ ou -1, et si $2n$ causes sont en jeu, une erreur résultante 2ε peut être obtenue conformément au théorème de BERNOULLI de $\binom{2n}{n+\varepsilon}$ manières différentes et la probabilité de l'erreur 2ε sera :

$$\binom{2n}{n+\varepsilon}\frac{1}{2^{2n}}.$$

Lorsqu'on mesure une grandeur inconnue z et M_i est la mesure obtenue, l'*erreur* commise ε_i, inconnue aussi, sera $\varepsilon_i = M_i - z$. On supposera que cette erreur est accidentelle, et que la probabilité pour qu'elle soit comprise entre $\varepsilon_i - \frac{1}{2}\Delta\varepsilon_i$ et $\varepsilon_i + \frac{1}{2}\Delta\varepsilon_i$ est donnée par

$$(1) \qquad P(\varepsilon_i)\Delta\varepsilon_i = \frac{1}{\zeta\sqrt{2\pi}}e^{-\varepsilon_i^2/2\zeta^2}\Delta\varepsilon_i;$$

c'est en même temps la probabilité pour que la mesure obtenue soit comprise entre $M_i - \frac{1}{2}\Delta M_i$ et $M_i + \frac{1}{2}\Delta M_i$, et on peut l'écrire :

$$P_1(M_i)\,\Delta M_i = \frac{1}{\zeta\sqrt{2\pi}}e^{-(M_i-z)^2/2\zeta^2}\Delta M_i.$$

C'est l'énoncé de *la loi des erreurs.* On en tire immédiatement la probabilité $\mathfrak{P}(\lambda)$ pour que la valeur absolue de l'erreur ε_i soit plus petite que λ ; on trouve

$$(2) \qquad \mathfrak{P}(\lambda) = \frac{2}{\sqrt{2\pi}}\int_0^{\lambda\zeta} e^{-\frac{1}{2}t^2}dt.$$

De la formule (1) on tire aussi celle de la probabilité des mesures simultanées $M_1, M_2, \ldots, M_n$, c'est-à-dire celle des erreurs simultanées $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$

$$(3) \qquad \mathfrak{P}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)\,\Delta\varepsilon_1 \ldots \Delta\varepsilon_n = \left|\frac{1}{\zeta\sqrt{2\pi}}\right|^n e^{-\Sigma\varepsilon_i^2/2\zeta^2}\,\Delta\varepsilon_1 \ldots \Delta\varepsilon_n.$$

A l'aide de (1) on peut montrer que ζ^2 est l'espérance mathématique du carré des erreurs c'est-à-dire

$$\zeta^2 = \int_{-\infty}^{\infty} \varepsilon_i^2 P(\varepsilon_i)\,d\varepsilon_i.$$

Lorsqu'on fait n mesures, et que n est grand, ζ^2 est égal d'une manière approchée à la moyenne des carrés des erreurs

$$\zeta^2 \sim \frac{1}{n}\Sigma(M_i - z)^2;$$

ζ^2 est appelé "*dispersion* des observations" et sa racine carrée ζ "*erreur quadratique* des observations".

On peut considérer ζ comme la mesure de la précision des observations. En effet lorsque dans deux séries d'observations les erreurs quadratiques sont ζ_1 et ζ_2, et qu'on trouve les probabilités égales pour que les erreurs soient respectivement plus petites que λ_1 et λ_2, alors de (2) il résulte:

$$\frac{\lambda_1}{\zeta_1} = \frac{\lambda_2}{\zeta_2}$$

donc les grandeurs λ_1 et λ_2 correspondant à la même probabilité sont proportionnelles aux erreurs quadratiques ζ_1 et ζ_2.

§ 2. Première remarque. La loi des erreurs n'est applicable qu'aux grandeurs M_i mesurées directement et non pas à une fonction quelconque $f(M_i)$ de ces mesures; autrement dit, l'erreur à considérer doit être la différence entre une mesure M_i et la grandeur à déterminer.

Il est vrai que certaines fonctions de M_i peuvent aussi suivre la loi des erreurs, mais cela doit être démontré dans chaque cas; de plus l'erreur quadratique de la fonction doit être déterminée en partant de celle des mesures. Cette remarque est utile, car il y a des auteurs qui appliquent la loi des erreurs à des fonctions de mesures sans s'inquiéter de la justifier.

Par exemple si l'on connait les distances d_i d'un foyer de tremblement de terre aux observatoires O (pour $i = 1, 2, \ldots, n$)

et qu'on observe le temps t_i qu'une onde sismique met pour y arriver, et qu'on cherche à déterminer la vitesse v de l'onde, comme ce sont les temps qu'on a mesurés, les erreurs à considérer sont les grandeurs $\frac{d_i}{v} - t_i$ et non pas $\frac{d_i}{t_i} - v$ comme on l'a fait quelquefois.

S e c o n d e r e m a r q u e. La difficulté en employant la formule (1) est que les grandeurs z et ζ sont incomnues. Pour remédier à cet inconvéniant, souvent on procède de la manière suivante : on fait n mesures et on détermine leur moyenne $\Sigma M_i/n$; puis on considère *l'écart* ξ_i entre la mesure M_i et la moyenne : $\xi_i = M_i - \frac{\Sigma M_i}{n}$, et finalement on remplace l'erreur ε_i par l'écart ξ_i, ce qui veut dire que l'on remplace z par la moyenne des mesures (on verra que c'est sa valeur la plus probable). Ce procédé ne conduit pas à des résultats satisfaisants. En effet si l'on désigne la moyenne des erreurs ε_i par η, alors de

$$\varepsilon_i = M_i - z$$

il suit

$$\eta = \frac{\Sigma M_i}{n} - z$$

donc η est aussi l'erreur de la moyenne des observations. De plus la définition de l'écart $\xi_i = M_i - \frac{\Sigma M_i}{n}$ conduit à $\Sigma \xi_i = 0$. Des équations précédentes on tire

$$\varepsilon_i = \xi_i + \eta \tag{4}$$

et par suite

$$\Sigma \varepsilon_i = n\eta.$$

En outre on a

$$\Sigma \varepsilon_i^2 = \Sigma \xi_i^2 + n\eta^2 = n(\sigma^2 + \eta^2)$$

ou, si n est grand

$$\zeta^2 \sim \sigma^2 + \eta^2.$$

Gauss (*Werke*, Tome IV, p. 6) dit que la somme des erreurs accidentelles devrait être nulle, et si l'on trouve que la moyenne des erreurs est égale à η, alors η forme la *partie constante de l'erreur*, c'est-à-dire la partie produite par des causes agissant toujours dans le même sens. En retranchant η de chaque observation il obtient des "observations corrigées". Les erreurs correspondant à ces observations sont évidemment égales aux écarts ξ_i. Cette manière de voir a conduit ses disciples à des conceptions erronées, en faisant croire qu'il était possible de découvrir, à l'aide de la loi des erreurs, les erreurs constantes ou systématiques, par exemple celles dues aux défauts des instruments de mesure, ce qui est impossible. D'autre part on verra que l'erreur η de la moyenne des mesures est aussi une erreur fortuite qui suit la loi des erreurs.

De plus, en remplaçant dans ζ la grandeur z par sa valeur la plus probable, on trouve $\zeta^2 = \sigma^2$. Enfin en écrivant $\zeta = \sigma$ dans la formule (3) qui donne la probabilité simultanée des mesures $M_1, \ldots, M_n$, on obtient

$$\left(\frac{1}{\sigma\sqrt{2\pi e}}\right)^n \Delta M_1 \ldots \Delta M_n$$

ce qui conduit à des résultats inacceptables. On en conclut qu'on ne peut remplacer simplement dans les formules la grandeur ζ par sa valeur la plus probable; on verra qu'il est préférable de déterminer la probabilité totale de z quel que soit ζ.

T r o i s i è m e r e m a r q u e. Pour déterminer la probabilité à posteriori de la grandeur inconnue z, on est obligé d'avoir recours au théorème de BAYES. A cet effet il faut adopter une hypothèse concernant les probabilités à priori de z (probabilités avant les mesures); généralement on considère à priori toutes les valeurs de z comme également probables, c'est-à-dire on admet que la probabilité pour la grandeur inconnue d'être comprisé entre $z - \frac{1}{2}\Delta z$ et $z + \frac{1}{2}\Delta z$ est proportionnelle à Δz. Cette hypothèse conduit a des conclusions très satisfaisantes. Il en résulte d'après le théorème de BAYES qu'après les n mesures, la probabilité à posteriori de z sera proportionnelle à (3), donc à

$$\left(\frac{1}{\zeta\sqrt{2\pi}}\right)^n e^{-\Sigma \varepsilon_i^2 2\zeta^2} \Delta\varepsilon_1 \ldots \Delta\varepsilon_n \Delta z.$$

Il faut remarquer, que dans cette formule ζ est une fonction de z; en effet on a, si n est grand, $\zeta^2 \sim \Sigma(M_i - z)^2/n$; mais les mesures M_i dépendent aussi de z, de plus la relation précédente n'est pas valable si n est petit; elle ne peut servir par exemple à déterminer la dérivée de ζ par rapport à z. Pour sortir de cette difficulté, le plus simple c'est de considérer la précision ζ des mesures dans une série d'observations indépendante des mesures M_i et de la grandeur z à mesurer. Comme ces quantités sont du même ordre de grandeur c'est très possible. Dans ces conditions on admettra que la probabilité à priori, c'est-à-dire avant les mesures, pour que la précision soit comprise entre $\zeta - \frac{1}{2}\Delta\zeta$ et $\zeta + \frac{1}{2}\Delta\zeta$ ne dépend que de ζ. A première vue on pourrait croire que toutes les valeurs de ζ sont également probables. Cette hypothèse conduirait aux résultats de GAUSS d'aprés lesquels la valeur la plus probable de ζ^2 est $\Sigma\xi_i^2/(n-1)$; mais en d'autres cas elle donnerait des résultats moins satisfaisants. Du reste, l'idée que la probabilité d'une erreur quadratique très grande, par exemple plus grande que la grandeur à mesurer soit aussi probable qu'une erreur quadratique petite, est contaire à notre manière de penser. [D'après la formule (2), la probabilité pour que la valeur absolue de l'erreur ε_i soit plus grande que ζ est 0,32. Il en résulte que si n est assez grand, environ un tiers des mesures dépassera $z + \zeta$, ce qui est certainement impossible si $\zeta > z$].

On en conclut qu'il vaut mieux adopter l'hypothèse que la probabilité de ζ diminue quand ζ augmente, et de prendre cette probabilité proportionnelle à $\Delta\zeta/\zeta^\alpha$. Pour $\alpha = 0$ cela donne l'hypothèse précédente. De plus on verra qu'il faut supposer α au moins égal à 2.

Quatrième remarque. GAUSS considère que la valeur la plus probable de ζ^2 est $n\sigma^2/(n-1)$, ce qui veut dire que la valeur la plus probable de la moyenne des carrés des erreurs est égale à la somme des carrés des écarts divisée par $n-1$, lorsqu'il s'agit d'une inconnue à déterminer, et divisé par $n-m$, lorsqu'il y en a m inconnues. (*Werke,* Tome IV, p. 49). L'idée est qu'étant données n équations et m inconnues ($n > m$) on peut considérer m équations comme satisfaites et $n-m$ équations donnant des erreurs; or même si l'on procédait ainsi, il faudrait tenir compte des erreurs nulles pour déterminer leur moyenne.

Cette remarque a une certaine importance, car la valeur adoptée pour ζ^2 influe sur les résultats numériques dans le calcul des moindres carrés. On verra que la valeur la plus probable de ζ^2 est $n\sigma^2/(n+\alpha-1)$; donc si on se contente d'une valeur approchée, $\zeta^2 = \sigma^2$ est préférable à la valeur $\zeta^2 = n\sigma^2/(n-1)$ proposée par GAUSS.

C i n q u i è m e r e m a r q u e. Quelquefois il faut faire entrer dans le calcul des mesures qui ne sont pas également précises. Pour y arriver il faut tenir compte de leur poids. On assigne à l'une des mesures, dont l'écart quadratique est ζ_1, *l'unité de poids* et on dit qu'une seconde mesure dont l'écart quadratique est ζ_2 a un poids p, lorsqu'il faudrait combiner p observations de la première mesure pour que leur moyenne soit aussi exacte que la seconde mesure. Nous verrons au § 4 que l'erreur quadratique de cette moyenne est $\zeta_1/\sqrt{p}$. On en conclut que les poids sont inversement proportionnels aux carrés des erreurs quadratiques:

$$\frac{p}{1} = \frac{\zeta_1^2}{\zeta_2^2}.$$

On verra au § 9, un exemple où l'on a déterminé le poids des valeurs x_ν calculées, lorsque le poids des mesures M_i était considéré comme égale à l'unité.

On ne peut rien objecter à cette manière de procéder. Il arrive que les observateurs considèrent certaines mesures de la même grandeur comme plus exactes que les autres. Cette différence peut être due à ce que les mesures sont exécutées à l'aide d'appareils de précisions différentes, ou par des observateurs plus ou moins exercés, ou encore dans des circonstances différentes. Si dans ces cas, on désire assigner des poids différents aux mesures, il faut prendre des précautions, en tous cas il faut fixer les poids *avant l'exécution des observations.* Car si on les fixait après les mesures, comme cela a été fait plusieurs fois, par exemple en donnant à chaque mesure un poids d'autant plus petit que son écart de la moyenne est plus grand, on arriverait à des résultats en désaccord avec la loi des erreurs.

§ 3. La loi (1) des erreurs a été formulée pour la première fois par ADRAIN dans ses Research concerning the probabilities of errors which happen in making observations (*The Analyst* or *Mathematical Museum,* 1 (1808), pp. 93—109). Il en a tiré plusieurs applications, entre autres il a remarqué qu'en vertu de cette formule la valeur la plus probable d'une grandeur inconnue, mesurée plusieurs fois, est la moyenne arithmétique des valeurs observées; de plus que la position la plus probable d'un point déterminé dans l'espace est le centre de gravité des points observés.

Trois ans avant, Legendre est arrivé à la même conclusion dans ces *Nouvelles méthodes pour la détermination des orbites des comètes* (Paris, 1806). Il dit: Étant donné la fonction

$$Y = a_1 x_1 + a_2 x_2 + \ldots + a_m x_m$$

où les a_ν sont des coefficients disponibles, on mesure le système de grandeurs x_ν plusieurs fois; supposons que l'$i^{\text{ième}}$ mesure ait donné

$$Y_i = a_1 x_{1i} + a_2 x_{2i} + \ldots + a_m x_{mi}$$

correspondant à une erreur $Y_i - Y$. Généralement le nombre des mesures est très grand (c'est toujours le cas dans les problèmes physiques et astronomiques), comme le nombre des équations surpasse celui des inconnues, il est impossible de disposer de ces inconnues pour annuler toutes les erreurs. Il faudra donc se contenter de les réduire autant que possible. Pour y arriver, différentes considérations ont amené Legendre à disposer des inconnues de manière à rendre minimum la somme des carrés des erreurs $\Sigma(Y_i - Y)^2$. (C'est la méthode utilisée aujourd'hui.) Dans l'ouvrage cité ci-dessus il dit "La règle par laquelle on prend le milieu entre les résultats de différentes observations (p. 74) n'est qu'une conséquence très simple de notre méthode générale que nous appellerons *Méthode des moindres quarrés*". Plus loin il remarque que d'une manière plus générale, si plusieurs points sont donnés dans l'espace, la somme des carrés des distances de ces points à leur centre de gravité est minimum.

En 1809 Gauss dans sa *Theoria Motus Corporum Coelestium* déduisait indépendemment de ce qui précède la même loi, et plus tard en 1823 et 1826 il a publié dans les *Commentationes Societatis Scientiarum Gottingensis* sa Theoria Combinationis Observationum Erroribus Minimis Obnoxiae. Dans cette dernière publication, il a donné une méthode complète de l'application de cette loi.

Laplace a déduit la loi des erreurs en partant de la supposition que chaque erreur est le résultat de la combinaison d'un très grand nombre d'erreurs élémentaires (*Théorie des probabilités*, 1812, p. 304). Plus tard Young dans ses Remarks on the probabilities of error in physical observations (*Philosophical Transactions*, 1819) et Hagen (*Grundzüge der Wahrscheinlichkeitsrechnung*, Berlin, 1837) ont perfectionné cette méthode. Plus tard d'autres géomètres ont contribué à la théorie mathématique des erreurs, par exemple: CROFTON, TCHÉBICHEF, ENCKE, BIENAYMÉ, ELLIS, GLAISHER, LINDELÖF, STORY, CHARLIER, POINCARÉ.

Supposons que la grandeur z ait été mesurée n fois et que $M_1, M_2, \ldots, M_n$ soient les mesures obtenues. Désignons par $\varphi(\varepsilon_i)$ la probabilité de l'erreur $\varepsilon_i = x_i - z$. Si nous posons les conditions: 1. la probabilité de l'erreur ε est égale à la probabilité de l'erreur $-\varepsilon$, donc $\varphi(\varepsilon) = \varphi(-\varepsilon)$; 2. le maximum de $\prod_{i=1}^{n} \varphi(M_i - z)$ est atteint pour z égal à la moyenne des mesures M_i; ces conditions conduisent à la loi (1).

Pour se rendre compte si certaines erreurs suivent la loi du hasard, remarquons que dans ce cas, si le nombre des mesures est grand, le nombre des erreurs positives doit être à peu près égal au nombre des erreurs négatives; la somme des erreurs doit être sensiblement égale à zéro. La moyenne des carrés des écarts divisée par le carré de la moyenne des valeurs absolues des écarts doit être à peu prés $\frac{1}{2}\pi$.

Si la probabilité de l'erreur ε_i *suit la loi* (1), il faut que la probabilité totale de toutes les erreurs soit égale à l'unité. En réalité on a

$$\frac{1}{\zeta\sqrt{2\pi}} \int_{-z}^{\infty} e^{-\varepsilon^2/2\zeta^2} d\varepsilon = \frac{1}{\sqrt{2\pi}} \int_{-z/\zeta}^{\infty} e^{-\frac{1}{2}t^2} dt \sim 1,$$

en effet lorsque z est grand relativement à ζ, cette intégrale est à peu près égale à l'unité. On en conclut que dans ce cas les erreurs peuvent, d'une manière approchée, suivre la loi (1) du hasard. Par contre lorsque les erreurs sont du même ordre de grandeur que la quantité z à mesurer, la loi n'est pas applicable.

§ 4. Pour obtenir la probabilité $\mathfrak{P}(\lambda)$ que *la valeur absolue de l'erreur de la moyenne* des n mesures soit plus petite que λ, c'est-à-dire $|\eta| < \lambda$, il faut effectuer la sommation:

$$\mathfrak{P}(\lambda) = \Sigma \Delta P(M_1, \ldots, M_n)$$

pour toutes les valeurs de M_i satisfaisant à $|\Sigma\varepsilon_i| < n\lambda$.

Pour faciliter cette sommation considérons un facteur discontinu $F(n\lambda, \Sigma\varepsilon_\nu)$ tel que l'on ait $F = 1$ lorsque $|\Sigma\varepsilon_\nu| < n\lambda$ et $F = 0$ lorsque $|\Sigma\varepsilon_\nu| > n\lambda$. Comme la probabilité pour avoir exactement

$|\Sigma\varepsilon_\nu| = n\lambda$ est nulle, il suffit que F ait une valeur finie dans ce cas. Il en résulte qu'après avoir multiplié l'expression (3) par F, on peut faire la somme par rapport aux M_i de zéro à ∞ pour toutes les valeurs de i.

$$F = \frac{2}{\pi}\int_0^\infty \sin(n\lambda u)\cos(u\Sigma\varepsilon_\nu)\frac{du}{u}$$

est un tel facteur.

En multipliant (3) par F on trouve

$$\mathfrak{P}(\lambda) = \frac{2}{\pi}\left(\frac{1}{\zeta\sqrt{2\pi}}\right)^n \cdot$$

$$\cdot\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} e^{-\Sigma\varepsilon_\nu^2/2\zeta^2}\int_0^\infty \sin(n\lambda u)\cos(u\Sigma\varepsilon_\nu)\frac{du}{u}d\varepsilon_1\ldots d\varepsilon_n.$$

En remarquant que $\cos(u\Sigma\varepsilon_\nu)$ est la partie réelle de $e^{iu\Sigma\varepsilon_\nu}$, on aura des intégrales de la forme

$$\int_{-\infty}^{\infty} e^{-\varepsilon_\nu^2/2\zeta^2 + iu\varepsilon_\nu}d\varepsilon_\nu = \zeta\sqrt{2\pi}\,e^{-\frac{1}{2}u^2\zeta^2}$$

pour toutes les valeurs de ν. On en conclut

$$\mathfrak{P}(\lambda) = \frac{2}{\pi}\int_0^\infty e^{-\frac{1}{2}nu^2\zeta^2}\sin(n\lambda u)\frac{du}{u}.$$

Posons encore, $\omega = \zeta u\sqrt{n}$ et $x = \lambda\sqrt{n}/\zeta$; la formule précédente deviendra

$$\mathfrak{P}(\lambda) = \frac{2}{\pi}\int_0^\infty e^{-\frac{1}{2}\omega^2}\sin(x\omega)\frac{d\omega}{\omega};$$

mais on sait que cette intégrale est égale à

$$\mathfrak{P}(\lambda) = 2\int_0^x \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}dt \tag{5}$$

où $x = \lambda\sqrt{n}/\zeta$. On en conclut que $\mathfrak{P}(\lambda)$, la probabilité pour que la valeur absolue de l'erreur de la *moyenne* soit plus petite que λ est en même temps égale à la probabilité pour que la valeur absolue de l'erreur d'une *seule observation* soit plus petite que $\lambda\sqrt{n}$;

de plus elle est encore égale à la probabilité pour que *la somme* de erreurs de n observations soit plus petite que λn. En outre on en conclut, que l'erreur de la moyenne des mesures suit la même loi que l'erreur d'une simple mesure; de plus que si ζ est l'erreur quadratique des mesures M_i, alors celle de la moyenne sera égale à $\zeta/\sqrt{n}$.

On en tire par exemple, que la probabilité est la même pour avoir

$$|\varepsilon_i| < 1 \quad \text{ou} \quad |\Sigma\varepsilon_i| < \sqrt{n} \quad \text{ou encore} \quad \left|\frac{\Sigma\varepsilon_i}{n}\right| < \frac{1}{\sqrt{n}}.$$

Fourier connaissait cette règle bien avant que la théorie des erreurs fût découverte: "Pour prendre la hauteur de la pyramide de Chéops, il fit simplement mesurer par des soldats les 203 marches de ce gigantesque escalier. Vos hommes manquent d'habitude, disait-on, les surfaces sont irrégulières, les arêtes sont inclinées, aucune précision n'est possible, et l'erreur commise sur chaque marche sera multiplié par 203. Elle le sera par 14 seulement, répondit-il résolument car 14 est la racine carrée de 203." (Bertrand, *Calcul des Probabilités,* p. XXXVIII.)

§ 5. Probabilité de l'erreur d'une fonction linéaire de m **grandeurs mesurées.** Soit

$$Y = a_1 z_1 + a_2 z_2 + \ldots + a_m z_m.$$

Les coefficients a_ν sont donnés et on connait les valeurs des z_ν; supposons que les mesures des z_ν aient fourni les grandeurs $M_{\nu i}$; de manière que l'$i^{\text{ième}}$ mesure donne

$$Y_i = a_1 M_{1i} + a_2 M_{2i} + \ldots + a_m M_{mi}. \tag{6}$$

L'erreur de Y_i sera

$$\bar{\varepsilon}_i = Y_i - Y = a_1(M_{1i} - z_i) + \ldots + a_m(M_{mi} - z_m) = a_1\varepsilon_{1i} + \ldots + a_m\varepsilon_{mi}.$$

On tire de l'équation (6) la moyenne des Y_i

$$\frac{\Sigma Y_i}{n} = a_1 \frac{\Sigma M_{1i}}{n} + \ldots + a_m \frac{\Sigma M_{mi}}{n}.$$

De plus l'écart $\bar{\xi}_i$ entre Y_i et sa moyenne sera

$$\bar{\xi}_i = Y_i - \frac{\Sigma Y_i}{n} = a_1\left[M_{1i} - \frac{\Sigma M_{1i}}{n}\right] + \ldots + a_m\left[M_{mi} - \frac{\Sigma M_{mi}}{n}\right]$$

ou encore

$$\bar{\xi}_i = a_1\xi_{1i} + \ldots + a_m\xi_{mi},$$

où les $\xi_{\nu i}$ sont les écarts respectifs.

Enfin $\bar{\eta}$, l'erreur de la moyenne des Y_i, sera

(7) $$\bar{\eta} = \frac{\Sigma Y_i}{n} - Y = a_1\left[\frac{\Sigma M_{1i}}{n} - z_1\right] + \ldots + a_m\left[\frac{\Sigma M_{mi}}{n} - z_m\right],$$

on trouve donc comme précédemment $\bar{\varepsilon}_i = \bar{\xi}_i + \bar{\eta}$.

La *probabilité d'obtenir les erreurs* $\varepsilon_{1i}, \ldots, \varepsilon_{mi}$ est tirée de la formule (1)

(8) $$\frac{1}{\zeta_1 \ldots \zeta_m (2\pi)^{\frac{1}{2}m}} e^{\sum_\nu \varepsilon_{\nu i}^2/2\zeta_\nu^2} \Delta\varepsilon_1 \ldots \Delta\varepsilon_m.$$

Pour obtenir la probabilité $\mathfrak{P}(\lambda)$ que *la valeur absolue de l'erreur de la moyenne* des Y_i soit plus petite que λ, il suffit de faire la somme de la probabilité (8) pour toutes les valeurs de $\varepsilon_{\nu i}$ telles que

$$|\bar{\eta}| = |a_{1i}\varepsilon_{1i} + \ldots + a_{mi}\varepsilon_{mi}| < \lambda.$$

On y arrive le plus simplement si on multiplie (8) par le facteur discontinu $F(\lambda, \Sigma a_\nu \varepsilon_{\nu i})$ qui est égal à l'unité pour les valeurs satisfaisant à l'inégalité précédente et nul pour les autres; alors on peut effectuer les sommations par rapport aux ε_ν de $-\infty$ à ∞; le même calcul que celui qu'on a employé pour déterminer la probabilité de l'erreur de la moyenne au § 4, conduit à

$$\mathfrak{P}(\lambda) = \frac{2}{\pi}\int_0^\infty e^{-\frac{1}{2}u^2} \sin u\lambda \frac{du}{u}.$$

Posons maintenant $\omega = u\sqrt{\Sigma a_\nu^2 \zeta_\nu^2}$ et $u\lambda = \omega x$; on obtient

$$\mathfrak{P}(\lambda) = \frac{2}{\pi}\int_0^\infty e^{-\frac{1}{2}\omega^2} \sin \omega x \frac{d\omega}{\omega},$$

où $x = \lambda/\sqrt{\Sigma a_\nu^2 \zeta_\nu^2} = \lambda/\bar{\zeta}$.

Par suite

(9) $$\bar{\zeta}^2 = \Sigma a_\nu^2 \zeta_\nu^2$$

est la moyenne des carrés des erreurs des Y_i. On en conclut que si les z_ν suivent la loi (1), la fonction Y la suivra aussi, de plus l'écart quadratique de Y sera $\bar{\zeta}$.

Cas particulier. Soit $Y = \frac{1}{m}(z_1 + z_2 + \ldots + z_m)$, autrement dit Y est la moyenne des quantités z_ν; en outre si l'on a $\zeta_\nu = \zeta$, alors de (9) il résulte que l'erreur quadratique de la moyenne des z_ν sera

$$\bar{\zeta} = \sqrt{m \frac{1}{m^2} \zeta^2} = \frac{\zeta}{\sqrt{m}}$$

conformément à ce qu'on a trouvé précédemment, mais cette fois m n'est pas nécessairement grand.

§ 6. Etant données une grandeur z et son erreur quadratique déterminée préalablement, on fait n mesures $M_1, M_2, \ldots, M_n$ et l'on demande la *probabilité pour que l'erreur quadratique* correspondant aux observations *soit plus petite que* λ. Comme $M_i - z = \varepsilon_i$ cette erreur quadratique est $\sqrt{\Sigma \varepsilon_i^2 / n}$; par suite, il faut déterminer la probabilité d'avoir

$$\Sigma \varepsilon_i^2 < n \lambda^2. \tag{10}$$

La probabilité d'obtenir les mesures M_i, ou ce qui revient au même, la probabilité des erreurs ε_i est donnée par

$$\left(\frac{1}{\zeta\sqrt{2\pi}}\right)^n e^{-\Sigma \varepsilon_i^2 / 2\zeta^2} \Delta\varepsilon_1 \ldots \Delta\varepsilon_n. \tag{11}$$

Pour obtenir la probabilité cherchée il faut faire la somme des probabilités (11) correspondant à chaque système ε_i satisfaisant à l'inégalité (10). On détermine cette somme, en faisant correspondre à chaque système de ε_i un point de l'espace à n dimensions. Les points favorables seront à l'intérieur de la sphère de rayon $R = \lambda\sqrt{n}$. La probabilité est constante à l'intérieur de la couche sphérique de rayon r et $r + dr$, où

$$r^2 = \varepsilon_1^2 + \varepsilon_2^2 + \ldots + \varepsilon_n^2.$$

Le volume de cette couche est égal à

$$\frac{2\pi^{\frac{1}{2}n} r^{n-1}}{\Gamma\left(\frac{1}{2}n\right)} dr.$$

La probabilité cherchée est proportionnelle à ce volume et à la probabilité (11). On a donc

$$d\mathfrak{P} = \frac{2C}{\Gamma\left(\frac{1}{2}n\right)} \left(\frac{1}{\zeta\sqrt{2}}\right)^n r^{n-1} e^{-r^2/2\zeta^2} dr. \tag{12}$$

Pour déterminer la valeur de C remarquons que l'intégrale de cette expression, r variant de zéro à ∞, doit être égale à l'unité; en effet c'est la probabilité pour que l'erreur quadratique soit quelconque. Pour effectuer l'intégration introduisons une nouvelle variable $\omega = r^2/2\zeta^2$ d'où $d\omega = r\,dr/\zeta^2$. On a

$$\int_0^\infty d\mathfrak{P} = \frac{C}{\Gamma\left(\frac{1}{2}n\right)} \int_0^\infty \omega^{\frac{1}{2}n-1}\, e^{-\omega}\, d\omega = C = 1.$$

Par suite dans la formule (12), la constante C est égale à l'unité.

Finalement dans le cas de la probabilité cherchée on doit avoir $r^2 < n\lambda^2$; par suite $\omega < n\lambda^2/2\zeta^2 = x$; et

$$\mathfrak{P}(n\lambda^2) = \frac{1}{\Gamma\left(\frac{1}{2}n\right)} \int_0^x \omega^{\frac{1}{2}n-1}\, e^{-\omega}\, d\omega = I(\bar{u}, \bar{p}) \tag{13}$$

où $I(\bar{u}, \bar{p})$ est le rapport de la fonction gamma incomplète $\Gamma_x\left(\frac{1}{2}n\right)$ à la fonction complète correspondante $\Gamma\left(\frac{1}{2}n\right)$. On a

$$\bar{p} = \frac{1}{2}n - 1 \qquad \text{et} \qquad \bar{u} = \frac{\lambda^2}{\zeta^2}\sqrt{\frac{1}{2}n}\,.$$

Cas particulier 1. La probabilité pour que l'on ait $\Sigma\varepsilon_i^2 < n\zeta^2$ est donnée par la formule (13)

$$\mathfrak{P}(n\zeta^2) = I\left(\sqrt{\frac{1}{2}n},\ \frac{1}{2}n - 1\right).$$

Cette probabilité est d'après les tables de la fonction gamma incomplète un peu plus grand qu'une demi.

Cas particulier 2. En posant $n = 1$ dans la formule (13) on obtient la probabilité pour que le carré de l'erreur d'une observation soit plus petit que λ^2. On trouve

$$\frac{1}{\sqrt{\pi}} \int_0^{\lambda^2/2\zeta^2} \omega^{-\frac{1}{2}}\, e^{-\omega}\, d\omega = I\left(\frac{\lambda^2}{\zeta^2\sqrt{2}},\ -\frac{1}{2}\right),$$

c'est aussi la probabilité pour que la valeur absolue de l'erreur d'une mesure soit plus petite que λ; elle est donnée par la formule de LAPLACE:

$$\frac{2}{\sqrt{2\pi}} \int_0^{\lambda/\zeta} e^{-\frac{1}{2}t^2}\, dt.$$

On peut donc exprimer ce résultat par une fonction gamma-incomplète

$$I\left(\frac{\lambda^2}{\zeta^2\sqrt{2}}, -\frac{1}{2}\right) = \frac{2}{\sqrt{2\pi}} \int_0^{\lambda/\zeta} e^{-t^2/2}\,dt.$$

Remarque. L'intégration par parties de $I(\bar{u}, \bar{p})$ dans (13) donne lorsque $n = 2\nu$

$$I(\bar{u}, \bar{p}) = 1 - e^{-x} \sum_{i=0}^{\nu-1} \frac{x^{\nu-1-i}}{\Gamma(\nu-i)}$$

et lorsque $n = 2\nu + 1$

$$I(\bar{u}, \bar{p}) = I\left(x\sqrt{2}, -\frac{1}{2}\right) - e^{-x} \sum_{i=0}^{\nu-1} \frac{x^{\nu-\frac{1}{2}-i}}{\Gamma\left(\nu + \frac{1}{2} - i\right)}.$$

Dans ces formules $\bar{u} = x\sqrt{2}/\sqrt{n}$ et $\bar{p} = \frac{1}{2}n - 1$.

Pour obtenir la probabilité que la *moyenne des carrés des erreurs* soit comprise entre λ^2 et $\lambda^2 + \Delta\lambda^2$ il suffit de poser dans (12) $r^2 = n\lambda^2$; on obtient

$$(14) \qquad \frac{n}{2\zeta^2\,\Gamma\left(\frac{1}{2}n\right)} \left(\frac{n\lambda^2}{2\zeta^2}\right)^{\frac{1}{2}n-1} e^{-n\lambda^2/2\zeta^2}\,d\lambda^2.$$

Cette probabilité est nulle pour $\lambda = 0$ et $\lambda = \infty$, elle est maximum pour $\lambda^2 = \frac{n-2}{n}\zeta^2$.

En multipliant la formule (14) par λ^2 et en intégrant de zéro à ∞, on a l'espérance mathématique $\mathcal{E}(\lambda^2)$ des grandeurs λ^2. La méthode appliquée dans le cas de la formule (12) conduit à $\mathcal{E}(\lambda^2) = \zeta^2$.

En multipliant la formule (14) par λ^4 et en intégrant on a l'espérance mathématique des λ^4, $\mathcal{E}(\lambda^4) = \frac{n+2}{n}\zeta^4$.

Finalement le carré de l'erreur quadratique de λ^2 sera

$$\mathcal{E}(\lambda^2 - \zeta^2)^2 = \mathcal{E}(\lambda^4) - \mathcal{E}^2(\lambda^2) = \frac{2}{n}\zeta^4;$$

enfin l'erreur quadratique de λ^2 est $\zeta^2\sqrt{\frac{2}{n}}$.

§ 7. Si z était connu et ζ inconnu on admettra que la probabilité à priori de ζ est proportionnelle à $\Delta\zeta/\zeta^\alpha$. Après avoir fait n mesures, $\Delta V_1(\zeta)$, la probabilité à posteriori de ζ sera donnée par le théorème de BAYES. En posant dans (3)

$$s^2 = \Sigma\varepsilon_i^2 = n\eta^2 + \Sigma\xi_i^2 = n(\eta^2+\sigma^2)$$

on aura

$$\Delta V_1(\zeta) = \frac{\zeta^{-n-\alpha}e^{-s^2/2\zeta^2}\Delta\zeta}{\int\limits_0^\infty \zeta^{-n-\alpha}e^{-s^2/2\zeta^2}d\zeta}.$$

Transformons cette intégrale en posant $\omega = s^2/2\zeta^2$ et par suite $d\omega = -s^2 d\zeta/\zeta^3$; le dénominateur devient:

$$(15)\qquad \frac{2^{\frac{1}{2}(n+\alpha-3)}}{s^{n-1+\alpha}}\int\limits_0^\infty \omega^{\frac{1}{2}(n+\alpha-3)}e^{-\omega}d\omega = \frac{2^{\frac{1}{2}(n+\alpha-3)}}{s^{n+\alpha-1}}\Gamma\left(\frac{n+\alpha-1}{2}\right).$$

Enfin la *probabilité à posteriori de* ζ sera si z est connu:

$$\Delta V_1(\zeta) = \frac{\omega^{\frac{1}{2}(n+\alpha-3)}e^{-\omega}\Delta\omega}{\Gamma\left(\frac{n+\alpha-1}{2}\right)} = \frac{s^{n+\alpha-1}e^{-s^2/2\zeta^2}\Delta\zeta}{2^{\frac{1}{2}(n+\alpha-3)}\Gamma\left(\frac{n+\alpha-1}{2}\right)\zeta^{n+\alpha}}.$$

Cette probabilité sera maximum pour $\zeta^2 = \dfrac{n}{n+\alpha}(\eta^2+\sigma^2) = \dfrac{\Sigma\varepsilon_i^2}{n+\alpha}$.

§ 8. Lorsque z est inconnu, mais que la précision ζ est connue, ayant été déterminée par des observations antérieures, alors la probabilité (1) sera maximum quand la somme des carrés des erreurs $\Sigma\varepsilon_i^2 = \Sigma(M_i - z)^2$ est minimum. On en conclut que la loi d'ADRAIN (1) conduit aussi au principe des moindres carrés.

Lorsque z et ζ sont inconnus et que l'on admet que leur probabilités à priori sont respectivement proportionnelles à Δz et $\Delta\zeta/\zeta^\alpha$ alors, après les n mesures, la probabilité à posteriori de z et ζ sera

$$(16)\qquad \Delta V_2(z,\zeta) = \frac{\zeta^{-n-\alpha}e^{-\Sigma(M_i-z)^2/2\zeta^2}\Delta z\,\Delta\zeta}{\int\limits_{-\infty}^{\infty}\int\limits_0^\infty \zeta^{-n-\alpha}e^{-\Sigma(M_i-z)^2/2\zeta^2}dz\,d\zeta}.$$

Pour obtenir la probabilité de z, on remplace quelquefois dans cette formule ζ par sa valeur la plus probable, mais ce procédé ne conduit pas à des résultats satisfaisants. Il vaut mieux déterminer la probabilité totale de z, quel que soit ζ; il suffit pour cela d'intégrer le numérateur et le dénominateur par rapport à ζ. En vertu de (15) l'intégration du numérateur donnera (comme $\Delta z = \Delta \eta$)

$$\frac{2^{\frac{1}{2}(n+\alpha-3)}\Gamma\left(\frac{n+\alpha-1}{2}\right)}{[n(\sigma^2+\eta^2)]^{\frac{1}{2}(n+\alpha-1)}}\Delta\eta;$$

il en résulte

$$\Delta V(z) = \frac{(\sigma^2+\eta^2)^{-\frac{1}{2}(n+\alpha-1)}\Delta\eta}{\int\limits_0^\infty (\sigma^2+\eta^2)^{-\frac{1}{2}(n+\alpha-1)}\,d\eta}.$$

Posons $\tan\varphi = \eta/\sigma$ et par suite $d\eta = \sigma d\varphi/\cos^2\varphi$. L'intégrale du dénominateur sera :

$$\sigma^{-n-\alpha+2}\int\limits_0^{1/2\,\pi} \cos^{n+\alpha-3}\varphi\,d\varphi = \frac{\pi\Gamma(n+\alpha-2)}{(2\sigma)^{n+\alpha-2}\Gamma^2\left(\frac{n+\alpha-1}{2}\right)}.$$

(Voir Bierens de Haan, Table 41, formule 3). Par suite la probabilité à posteriori de z et en même temps *la probabilité à posteriori de l'erreur η de la moyenne* des mesures sera quelle que soit la précision inconnue ζ

$$\Delta V(z) = \frac{\Gamma^2\left(\frac{n+\alpha-1}{2}\right)}{\Gamma(n+\alpha-2)}\cdot\frac{1}{2\pi\sigma}\left(\frac{4\sigma^2}{\sigma^2+\eta^2}\right)^{\frac{1}{2}(n+\alpha-1)}\Delta\eta$$

ou encore

$$\Delta V(z) = \frac{2^{n+\alpha-2}\Gamma^2\left(\frac{n+\alpha-1}{2}\right)}{\pi\Gamma(n+\alpha-2)}\cos^{n+\alpha-3}\varphi\,d\varphi. \tag{17}$$

Cette probabilité est maximum lorsque $\eta=0$ ou $z=\Sigma M_i/n$; mais alors z rend minimum la somme des carrés des erreurs; ainsi le principe des moindres carrés est valable dans ce cas aussi.

De la formule (17) il suit immédiatement la probabilité $\mathfrak{P}$ pour que la *valeur absolue de l'erreur de la moyenne* soit plus petite que λ, c'est-à-dire $|\eta|<\lambda$. Comme $\eta=\sigma\tan\varphi$ et $\lambda=\sigma\tan\varphi_1$, il faut avoir $\varphi<\varphi_1$, où $\varphi_1=\arctan\lambda/\sigma$. On trouve quelle que soit la préciison ζ

$$\mathfrak{P} = \frac{2^{n+\alpha-2}\Gamma^2\left(\frac{n+\alpha-1}{2}\right)}{\pi\Gamma(n+\alpha-2)}\int\limits_0^{\varphi_1}\cos^{n+\alpha-3}\varphi\,d\varphi.$$

Exemple 1. $n=4$, $\alpha=0$, on a

$$\vartheta=\sin\varphi_1=\sqrt{\frac{\lambda^2}{\sigma^2+\lambda^2}}$$

et si $\lambda=\sigma$ alors $\vartheta=1/\sqrt{2}$.

Pour obtenir des valeurs acceptables depuis $n=1$ il faut poser au moins $\alpha=2$, alors on trouve pour $\lambda=\sigma$:

$n=1$	$n=2$	$n=3$
$\vartheta=\frac{1}{2}$	$\vartheta=1/\sqrt{2}$	$\vartheta=\frac{1}{\pi}+\frac{1}{2}$.

De la même manière on peut déterminer la probabilité de ζ quel que soit z, en intégrant le numérateur et le dénominateur de (16) par rapport à z. Le numérateur sera:

$$\zeta^{-n-\alpha}e^{-n\sigma^2/2\zeta^2}\Delta\zeta\int_{-\infty}^{\infty}e^{-n\eta^2/2\zeta^2}d\eta,$$

en posant $t=\eta\sqrt{n}/\zeta$ et par suite $d\eta=\zeta dt/\sqrt{n}$, on obtient

$$\frac{\sqrt{2\pi}\,e^{-n\sigma^2/2\zeta^2}\Delta\zeta}{\zeta^{n+\alpha-1}\sqrt{n}}$$

Comme le dénominateur donne une valeur semblable on aura

$$V(\zeta)=\frac{\zeta^{-n-\alpha+1}e^{-n\sigma^2/2\zeta^2}\Delta\zeta}{\int_0^{\infty}\zeta^{-n-\alpha+1}e^{-n\sigma^2/2\zeta^2}d\zeta}.$$

Or l'intégrale du dénominateur est d'après (15) égale à

$$2^{\frac{1}{2}(n+\alpha-2)}\Gamma\left(\frac{n+\alpha-1}{2}\right)\Big/(n\sigma^2)^{\frac{1}{2}(n+\alpha-2)};$$

il s'ensuit la probabilité à posteriori de ζ quelle que soit la valeur inconnue de z:

$$V(\zeta)=\frac{(n\sigma^2)^{\frac{1}{2}(n+\alpha-2)}e^{-n\sigma^2/2\zeta^2}}{2^{\frac{1}{2}(n+\alpha-2)}\Gamma\left(\frac{n+\alpha-1}{2}\right)\zeta^{n+\alpha-1}}\Delta\zeta.$$

Le maximum de cette probabilité a lieu pour $\zeta^2=\frac{n\sigma^2}{n+\alpha-1}$. En posant dans cette formule $\alpha=0$, la valeur la plus probable de ζ^2 serait $\Sigma\varepsilon_i^2/(n-1)$ conformément au résultat proposé par Gauss.

§ 9. Dans le problème du § 5 les coefficients $a_1, a_2, \ldots, a_m$ étaient donnés et on a mesuré les grandeurs $z_1, z_2, \ldots, z_m$, puis on a cherché à déterminer les probabilités de $Y = a_1 z_1 + a_2 z_2 + \ldots + a_m z_m$. Il y a une autre catégorie de problèmes. Étant donnés les coefficients $a_{i1}, \ldots, a_{im}$ pour les valeurs de $i = 1, 2, \ldots, n$, on mesure les grandeurs suivantes

$$Y_i = \sum_{\nu=1}^{m} a_{i\nu} x_\nu .$$

Les quantités $x_1, x_2, \ldots, x_m$ sont inconnues et on demande leur probabilité. Soit M_i la mesure de Y_i, l'erreur inconnue sera:

$$\varepsilon_i = M_i - Y_i .$$

En supposant que les Y_i sont du même ordre de grandeur, et que la précision des observations est la même pour toutes les mesures l'erreur quadratique inconnue étant ζ, la probabilité à posteriori des grandeurs Y_i et par suite aussi des erreurs ε_i sera:

$$\varDelta P_1(\varepsilon_i, \ldots, \varepsilon_n, \zeta) = \frac{\zeta^{-n-\alpha} e^{-\Sigma \varepsilon_i^2/2\zeta^2} \varDelta \varepsilon_i \ldots \varDelta \varepsilon_n \varDelta \zeta}{\int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} \int\limits_{0}^{\infty} \zeta^{-n-\alpha} e^{-\Sigma \varepsilon_i^2/2\zeta^2} d\varepsilon_1 \ldots d\varepsilon_n d\zeta} .$$

En effectuant l'intégration au numérateur et au dénominateur par rapport à ζ comme dans le cas de la formule (16) on a *la probabilité à posteriori des Y_i* quelle que soit la précision ζ

$$(18) \qquad \varDelta P(\varepsilon_1, \ldots, \varepsilon_n) = \frac{(\Sigma \varepsilon_i^2)^{-\frac{1}{2}(n+\alpha-1)} \varDelta \varepsilon_1 \ldots \varDelta \varepsilon_n}{\int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} (\Sigma \varepsilon_i^2)^{-\frac{1}{2}(n+\alpha-1)} d\varepsilon_1 \ldots d\varepsilon_n} .$$

Cette probabilité sera encore maximum lorsque $\Sigma \varepsilon_i^2$ est minimum. Il s'agit donc de rendre minimum la quantité:

$$(19) \qquad \Sigma \varepsilon_i^2 = \sum_i (M_i - a_{i1} x_1 - a_{i2} x_2 - \ldots - a_{im} x_m)^2 .$$

Les équations de condition du minimum, dites équations normales, seront:

$$(20) \qquad \frac{\partial \Sigma \varepsilon_i^2}{\partial x_\nu} = \sum_i a_{i\nu} (M_i - a_{i1} x_1 - a_{i2} x_2 - \ldots - a_{im} x_m) = 0$$

pour $\nu = 1, 2, \ldots, m$. La somme des carrés des erreurs minima est

$$(21) \qquad \sum_i M_i (M_i - a_{i1} x_1 - a_{i2} x_2 - \ldots - a_{im} x_m) = \omega^2 ;$$

en effet en multipliant les équations (20) par x_ν et en retranchan les produits de l'équation précédente, on obtient l'équation (19). La valeur minimum ω^2 de $\Sigma\varepsilon_i^2$ est importante, car la probabilité pour que $\Sigma\varepsilon_i^2$ soit plus petite que ce minimum est nulle. Il faudra en tenir compte en effectuant les intégrations dans la formule (18).

Pour déterminer ω posons :

$$M_i = -a_{i,\,m+1}$$

et

$$x_{m+1} = 1\,;$$

de plus

$$\sum_{i=1}^{n} a_{i\nu} a_{is} = b_{\nu s}$$

donc

$$b_{\nu s} = b_{s\nu}.$$

Soit

$$D_i = \begin{vmatrix} b_{11} & b_{12} & \dots & b_{1i} \\ b_{21} & b_{22} & \dots & b_{2i} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \dots & b_{ii} \end{vmatrix},$$

alors les m équations normales s'écriront

$$b_{1\nu} x_1 + b_{2\nu} x_2 + \dots + b_{m+1,\,\nu} x_{m+1} = 0$$

pour $\nu = 1, 2, \dots, m$.

En outre l'équation (21) devient :

$$b_{1,\,m+1} x_1 + b_{2,\,m+1} x_2 + \dots + b_{m+1,\,m+1} x_{m+1} = \omega^2.$$

La résolution de ces $m+1$ équations à $m+1$ inconnues $x_1, \dots, x_{m+1}$ conduit à

$$x_{m+1} = \omega^2 D_m / D_{m+1}\,;$$

mais comme $x_{m+1} = 1$, on a $\omega^2 = D_{m+1}/D_m$.

Les premières m équations donnent, si l'on pose $x_{m+1} = 1$,

$$x_\nu = \frac{1}{D_m} \sum_{s=1}^{m} (-1)^{\nu+s+1} b_{m+1,\,s}\, \beta_{\nu s} \tag{22}$$

où $\beta_{\nu s}$ est le mineur de D_m obtenu en supprimant la ligne ν et la colonne s dans ce déterminant.

En partant de (22) on peut déterminer l'erreur quadratique ζ_ν de x_ν lorsqu'on connait l'erreur quadratique ζ des mesures M. En effet en rempaçant $b_{m+1,\,s}$ par sa valeur obtenue ci-dessus on trouve

$$x_\nu = \sum_{i=1}^{n} M_i \sum_{s=1}^{m} \frac{(-1)^{s+\nu}}{D_m} \beta_{\nu s} a_{is}$$

et en vertu de (9) le carré de l'erreur quadratique (ou la dispersion) de x_ν sera

$$\zeta_\nu^2 = \zeta^2 \sum_i \sum_s \sum_k \frac{(-1)^{s+k}}{D_m^2} \beta_{\nu s} \beta_{\nu k} a_{is} a_{ik};$$

mais

$$\sum_k (-1)^{s+k} \beta_{\nu k} b_{sk}$$

est égal à zéro si $s \neq \nu$ et égal à D_m si $s = \nu$; on en conclut que le carré de l'erreur quadratique de x_ν est

$$\zeta_\nu^2 = \frac{\beta_{\nu\nu}}{D_m} \zeta^2.$$

Lorsqu'on attribue aux mesures M_i l'unité de *poids*, alors le poids de la valeur x_ν obtenue (22), étant inversement proportionnelle aux dispersions respectives, sera

$$\frac{\zeta^2}{\zeta_\nu^2} = \frac{D_m}{\beta_{\nu\nu}}.$$

Pour intégrer l'équation (18) remarquons de nouveau que $\Sigma \varepsilon_i^2 = r^2$ est constant sur la surface de la sphère de rayon r dans l'espace à n dimensions; cette surface ayant pour aire $2\pi^{\frac{1}{2}n} r^{n-1} \Big/ \Gamma\left(\frac{n}{2}\right)$, donc l'intégrale du dénominateur de (18) devient

$$\int_\omega^\infty \frac{2\pi^{\frac{1}{2}n}}{\Gamma\left(\frac{n}{2}\right)} \frac{dr}{r^\alpha} = \frac{2\pi^{\frac{1}{2}n}}{\Gamma\left(\frac{n}{2}\right)} \frac{1}{(\alpha-1)\,\omega^{\alpha-1}},$$

il en résulte que la probabilité (18) est égale à

$$(23) \quad \Delta P(\varepsilon_1, \ldots, \varepsilon_n) = \\ = \frac{(\alpha-1)\,\Gamma\left(\frac{n}{2}\right)}{2\pi^{\frac{1}{2}n}} \left(\frac{D_{m+1}}{D_m}\right)^{\frac{1}{2}(\alpha-1)} (\Sigma \varepsilon_i^2)^{-\frac{1}{2}(n+\alpha-1)} \Delta\varepsilon_1 \ldots \Delta\varepsilon_n$$

et si l'on transforme le numérateur de (18) de la même manière que le dénominateur on trouve

$$\Delta P(\varepsilon_1, \ldots, \varepsilon_n) = (\alpha-1) \left(\frac{D_{m+1}}{D_m}\right)^{\frac{1}{2}(\alpha-1)} \frac{dr}{r^\alpha}.$$

Cette formule donne immédiatement *la probabilité pour que la somme des carrés des erreurs* des Y_i soit plus petite que λ^2 c'est-à-dire $\Sigma\varepsilon_i^2 = r^2 < \lambda^2$, quelle que soit la précision inconnue ζ. Il suffit d'intégrer l'expression précédente de zéro à λ, on trouve

$$\mathcal{S}(\lambda^2) = 1 - \left(\frac{D_{m+1}}{\lambda^2 D_m}\right)^{\frac{1}{2}(\alpha-1)}.$$

Ce résultat a été obtenu en supposant que la probabilité à priori de la précision ζ était proportionnelle à $\Delta\zeta/\zeta^\alpha$. Remarquons que pour $\alpha = 0$ (Gauss) la probabilité $\mathcal{S}(\lambda^2)$ diminuerait si λ^2 augmente, ce qui est impossible; de plus pour $\alpha = 1$ on aurait $\mathcal{S}(\lambda^2) = 1$ quel que soit λ, c'est inadmissible aussi. On en conclut comme cela a été remarqué précédemment que dans la probabilité à-priori de ζ il faut poser au moins $\alpha = 2$.

Lorsqu'on a mesuré n fois une grandeur inconnne, ne connaissant pas la précision ζ^2 des mesures, mais sachant que sa valeur la plus probable est (voir la fin du § 8)

$$\frac{\Sigma\varepsilon_i^2}{n+\alpha-1} = \frac{n\sigma^2}{n+\alpha-1},$$

on trouve dans l'hypothèse de Gauss correspondant à $\alpha = 0$ que cette valeur sera supérieure à σ^2 qui est le carré de l'écart quadratique des mesures. Dans l'hypothèse $\alpha = 1$ elle sera égale à σ^2 et enfin dans l'hypothèse $\alpha \geqq 2$, elle sera inférieur à σ^2. Nous avons vue que c'est cette dernière hypothèse qui est la mieux justifiée. Il en résulte que si l'on se contente d'une approximation pour la valeur la plus probable de la précision ζ^2, on peut prendre $\zeta^2 = \sigma^2$; mais comme c'est déja plus grande que la valeur rigoureuse, on n'a aucune raison d'employer celle proposée par Gauss, qui est encore plus grande.

(Reçu le 18 octobre 1941)

Zur Gaußischen Theorie der Reduktion binärer quadratischer Formen.

Von L. Rédei in Szeged.

In der Theorie der Äquivalenz der binären quadratischen Formen ist wohl der schwierigste Punkt der Nachweis, daß im Fall einer positiven Determinante äquivalente reduzierte Formen derselben Periode angehören. Aus diesem Satz folgt, daß die Klassenzahl gleich der Anzahl der Perioden ist. Eine nicht weniger wichtige Folgerung ist, daß die Perioden die Fundamentallösung der Pellschen Gleichung liefern; bis heute gibt es keine einfachere Methode zum Auffinden aller Lösungen dieser Gleichung.

Gausz[1]) hat den genannten Satz mit Hilfe von Kettenbrüchen bewiesen. Auf dieses fremde Mittel haben auch die späteren Autoren nicht verzichtet, obwohl eine Bestrebung vorlag, die Theorie eben in diesem Punkt zu vereinfachen. Im Gegenteil hat Dirichlet[2]) eine Vereinfachung unter anderem dadurch erreicht, daß er die Theorie der Kettenbrüche noch mehr zur Geltung brachte. Man sehe auch die Lehrbücher von Cahen[3]) und Dickson[4]).

Hier möchte ich einen Beweis mitteilen, der frei von Kettenbrüchen und auch sonst einfacher ist, als die bisherigen Beweise. Kurz werde ich auch darauf hinweisen, daß man dann auch bei der Anwendung auf die Pellsche Gleichung ohne Kettenbrüche auskommt.

---

1) C. F. Gausz, *Werke*, I (Leipzig, 1870), S. 180.

2) L. Dirichlet—R. Dedekind, *Vorlesungen über Zahlentheorie*, 4. Aufl. (Braunschweig, 1894), S. 199.

3) E. Cahen, *Théorie des nombres*, II (Paris, 1924), p. 305.

4) L. E. Dickson—E. Bodewig, *Einführung in die Zahlentheorie* (Leipzig und Berlin, 1931), S. 103.

Die Kenntnis des einschlägigen Kapitels im angeführten Lehrbuch von DIRICHLET setze ich voraus, übernehme auch die Bezeichnung fast unverändert, fühle mich trotzdem gezwungen, bequemlichkeitshalber einiges vorangeschickt zu wiederholen.

Unter einer binären quadratischen Form (kurz Form) verstehen wir mit GAUSZ

$$\varphi = ax^2 + 2bxy + cy^2$$

mit ganzen rationalen Koeffizienten a, b, c.[5]) Die Determinante $D = b^2 - ac$ soll stets positiv und keine Quadratzahl sein. Wir nennen

$$S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \qquad (\alpha\delta - \beta\gamma = 1)$$

mit rationalen $\alpha, \beta, \gamma, \delta$ eine Substitution. Unter $S\varphi$ werde die Form $ax'^2 + 2bx'y' + cy'^2$ verstanden, wobei $x' = \alpha x + \beta y$, $y' = \gamma x + \delta y$ ist. Gilt eine Beziehung $\psi = S\varphi$, so heißen φ und ψ äquivalent. Die Substitutionen

$$E = \begin{pmatrix} 0 & 1 \\ -1 & \delta \end{pmatrix}$$

heißen elementar, die $E\varphi$ die Nachbarformen von φ. Es gibt nur ein E, die φ in eine gegebene Nachbarform überführt.

Erste bzw. zweite Wurzel von φ sind durch

$$\omega = \frac{-b - \sqrt{D}}{c}, \quad \omega' = \frac{-b + \sqrt{D}}{c}$$

definiert, wobei wir $\sqrt{D}$ positiv annehmen. D und ω bestimmen die Form φ eindeutig. Sind $\overline{\omega}, \overline{\omega}'$ die entsprechenden Wurzeln einer zweiten Form $\overline{\varphi}$, so sind

$$\text{(1)} \qquad \overline{\varphi} = S\varphi, \quad \omega = \frac{\gamma + \delta\overline{\omega}}{\alpha + \beta\overline{\omega}}, \quad \omega' = \frac{\gamma + \delta\overline{\omega}'}{\alpha + \beta\overline{\omega}'}$$

gleiche Aussagen. Insbesondere sind φ und $\overline{\varphi}$ Nachbarformen dann und nur dann, wenn

$$\text{(2)} \qquad \omega + \frac{1}{\overline{\omega}} = r \qquad (r \text{ ganz rational})$$

ist, und zwar ist dann $\overline{\varphi} = \begin{pmatrix} 0 & 1 \\ -1 & r \end{pmatrix} \varphi$.

[5]) Ich wollte mich auch in dieser Hinsicht an Dirichlets Lehrbuch anpassen, obwohl bekanntlich Lagranges Definition $\varphi = ax^2 + bxy + cy^2$ sich mit der Zeit für glücklicher erwies. Unsere Arbeit läßt sich aber ohne jede Mühe für Lagranges Definition umschreiben.

Reduziert heißt φ, wenn

(3) $$|\omega| > 1, \ |\omega'| < 1, \ \omega\omega' < 0$$

ist. Nachher werden nur noch reduzierte Formen betrachtet. Unter ihnen hat jedes φ eine einzige Nachbarform $\overline{\varphi}$, und das entsprechende E (nämlich das mit $E\varphi = \overline{\varphi}$) ist durch

(4) $$\delta = \{\omega\}$$

bestimmt, wobei $\{\omega\}$ die ganze rationale Zahl zwischen 0 und ω bedeutet, die zu ω am nächsten liegt. (Wie hier, so auch später wollen wir mit „zwischen u und v" u und v selbst ausschließen. (4) ist sinnvoll, da $|\omega| > 1$ ist.)

Geht man also von einer reduzierten Form $\varphi = \varphi_1$ aus, so erhält man durch fortgesetzte Bildung von reduzierten Nachbarformen eine eindeutig bestimmte unendliche Folge $\varphi_1, \varphi_2, \varphi_3, \ldots$, die wir die Kette von φ nennen. Stets soll E_i die elementare Substitution bezeichnen, für die $E\varphi_i = \varphi_{i+1}$ ist. Jede Kette wiederholt sich periodisch, d. h. es gibt ein ganzes m so, daß $\varphi_1, \varphi_2, \ldots, \varphi_m$ verschieden sind, dagegen $\varphi_i = \varphi_k$ für $i \equiv k \pmod{m}$ gilt. $\varphi_1, \varphi_2, \ldots, \varphi_m$ heißt die Periode von φ, m die Periodenzahl.

Wir definieren noch $-S = \begin{pmatrix} -\alpha & -\beta \\ -\gamma & -\delta \end{pmatrix}$. Es sollen also S und $-S$ als verschieden betrachtet werden, obzwar immer $S\varphi = (-S)\varphi$ ist. Wir beweisen den Satz:

Es sei $\mathfrak{S} = (S, \varphi, \overline{\varphi})$ ein „System" bestehend aus einer Substitution $S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ und zwei reduzierten Formen $\varphi, \overline{\varphi}$ mit $S\varphi = \overline{\varphi}$. Abgesehen vom trivialen Fall $S = \pm\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ gilt genau eine der Gleichungen

(5_1) $$\operatorname{sgn} \overline{\omega} = \operatorname{sgn} \gamma\delta,$$

(5_2) $$\operatorname{sgn} \omega = -\operatorname{sgn} \alpha\gamma,$$

wobei $\operatorname{sgn} z = \frac{z}{|z|}$ $(z \neq 0)$ *ist. Gilt* (5_1), *so ist* $S = \pm E_1 E_2 \ldots E_n$ *mit irgendeinem n, zugleich also* $\overline{\varphi} = \varphi_{n+1}$. *Gilt aber* ($5_2$), *so gilt für das „umgekehrte" System* $\mathfrak{S}^{-1} = (S^{-1}, \overline{\varphi}, \varphi)$ *der* (5_1) *entsprechende Zusammenhang* (und also mit entsprechender Bezeichnung $S^{-1} = \pm \overline{E}_1 \overline{E}_2 \ldots \overline{E}_n$, $\varphi = \overline{\varphi}_{n+1}$).

Zum Beweis setzen wir

(6) $$|S| = |\alpha| + |\beta| + |\gamma| + |\delta|.$$

Da für $S = \pm\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ nichts behauptet wird, schließen wir diesen Fall aus.

Weiter zeigen wir, daß dann $\beta\gamma = 0$ unmöglich ist. Ist nämlich $\beta = 0$, so ist $\alpha\delta = 1$, $\gamma \neq 0$, und also nach (1) $\omega = \alpha\gamma + \overline{\omega}$, $\omega' = \alpha\gamma + \overline{\omega}'$. Aus dem letzteren folgt nach (3) $\alpha\gamma = \pm 1$, $\operatorname{sgn}\omega' = -\operatorname{sgn}\overline{\omega}'$, also auch $\operatorname{sgn}\omega = -\operatorname{sgn}\overline{\omega}$. Wieder nach (3) ist dann $|\omega - \overline{\omega}| > 2$, und das ist wegen $\omega - \overline{\omega} = \alpha\gamma$ unmöglich. Nehmen wir jetzt $\gamma = 0$ an. Dann folgt $\alpha\delta = 1$, $\beta \neq 0$, und nach (1) $\frac{1}{\omega} = \beta\delta + \frac{1}{\overline{\omega}}$, $\frac{1}{\omega'} = \beta\delta + \frac{1}{\overline{\omega}'}$. Jetzt folgt nach dem vorletzten Zusammenhang und (3), daß $\beta\delta = \pm 1$, $\operatorname{sgn}\omega = -\operatorname{sgn}\overline{\omega}$, also auch $\operatorname{sgn}\omega' = -\operatorname{sgn}\overline{\omega}'$ ist. Nach (3) heißt das $\left|\frac{1}{\omega'} - \frac{1}{\overline{\omega}'}\right| > 2$, und so ist mit $\frac{1}{\omega'} - \frac{1}{\overline{\omega}'} = \beta\delta$ wieder ein Widerspruch entstanden.

Auch noch den Fall $\alpha\delta = 0$ schicken wir voran. Es werde zuerst $\alpha = 0$. Dann ist $\beta\gamma = -1$, $S = \pm\begin{pmatrix}0 & 1\\ -1 & r\end{pmatrix}$ (r ganz rational). Da also $\pm S$ elementar ist, muß $\overline{\varphi}$ wegen $S\varphi = \overline{\varphi}$ gleich φ_2 und $S = \pm E_1$ sein. Nach (4) und (2) folgt hieraus $r = \{\omega\}$, $\operatorname{sgn}\overline{\omega} = -\operatorname{sgn} r$. Das bedeutet eben, daß jetzt (5_1) gilt. Endlich gilt (5_2) wegen $\alpha = 0$ nicht, und so ist der Satz für diesen Fall richtig. Es werde dann $\delta = 0$. Für $\mathfrak{S}^{-1} = \left(\begin{pmatrix}\delta & -\beta\\ -\gamma & \alpha\end{pmatrix}, \overline{\varphi}, \varphi\right)$ ist das der vorige Fall. Nach dem eben Gesehenen gilt (5_1) für dieses System, d. h. $\operatorname{sgn}\omega = -\operatorname{sgn}\alpha\gamma$, und das bedeutet eben, daß für $\mathfrak{S}$ (5_2) gilt. Dagegen gilt (5_1) für $\mathfrak{S}$ wegen $\delta = 0$ nicht, womit die Richtigkeit des Satzes auch für diesen Fall erwiesen ist.

Es ist nur noch der Fall $\alpha\beta\gamma\delta \neq 0$ übrig. Offenbar gilt $\alpha\beta\gamma\delta > 0$. Zuerst zeigen wir, daß aus (5_1) und (5_2) genau des eine gilt. Gilt nämlich (5_1), d. h. $\operatorname{sgn}\overline{\omega} = \operatorname{sgn}\gamma\delta = \operatorname{sgn}\alpha\beta$, so ist $\operatorname{sgn}\delta\overline{\omega} = \operatorname{sgn}\gamma$, $\operatorname{sgn}\beta\overline{\omega} = \operatorname{sgn}\alpha$. Hieraus und aus (1) folgt $\operatorname{sgn}\omega = \operatorname{sgn}\frac{\gamma}{\alpha}$, d. h. (5_2) gilt nicht. Gilt dagegen (5_1) nicht, ist also $\operatorname{sgn}\overline{\omega} = -\operatorname{sgn}\gamma\delta$, so folgt $\operatorname{sgn}\overline{\omega}' = \operatorname{sgn}\gamma\delta$ und weiter hieraus ebenso wie vorher $\operatorname{sgn}\omega' = \operatorname{sgn}\frac{\gamma}{\alpha}$. Dies bedeutet wirklich eben, daß (5_2) gilt.

Mit der letzten Behauptung des Satzes sind wir durch die Bemerkung fertig, daß (5_2) (für $\mathfrak{S}$) dasselbe ist wie (5_1) für $\mathfrak{S}^{-1}=\left(\begin{pmatrix}\delta & -\beta\\ -\gamma & \alpha\end{pmatrix}, \overline{\varphi}, \varphi\right)$.

Wir müssen nur noch den Hauptpunkt beweisen, daß (im Fall $\alpha\beta\gamma\delta\neq 0$) aus (5_1) $S=\pm E_1E_2\ldots E_n$ folgt. Das tun wir durch Induktion, indem wir die Behauptung für jeden ähnlichen Fall voraussetzen, wenn nur $|S|$ einen „kleineren" Wert hat. Wir setzen

$$r=\{\omega\}. \tag{7}$$

Nach (4) ist dann $E_1=\begin{pmatrix}0 & 1\\ -1 & r\end{pmatrix}$. Weiter setzen wir $T=E_1^{-1}S$. Wegen $\alpha\beta\gamma\delta\neq 0$ ist $T\neq\pm\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$. Es ist $T\varphi_2=T(E_1\varphi)=S\varphi=\overline{\varphi}$, und das heißt, daß auch $\mathfrak{T}=(T, \varphi_2, \overline{\varphi})$ ein System ist. Zeigen wir, daß auch für dieses Fall (5_1) vorliegt und $|T|<|S|$ ist, so folgt hieraus $T=\pm E_2E_3\ldots E_n$ $(n\geqq 2)$ entweder nämlich nach der Voraussetzung oder nach dem schon bewiesenen Fall $\alpha\beta\gamma\delta=0$ des Satzes. Dann folgt weiter $S=E_1T=\pm E_1E_2\ldots E_n$, und so werden wir den Satz auf diesem Wege bewiesen haben.

Nun ist

$$T=\begin{pmatrix}r & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix},$$

also

$$T=\begin{pmatrix}\gamma' & \delta'\\ \alpha & \beta\end{pmatrix} \text{ mit } \gamma'=\alpha r-\gamma,\ \delta'=\beta r-\delta. \tag{8}$$

Aus (5_1) folgt aber $\operatorname{sgn}\overline{\omega}=\operatorname{sgn}\alpha\beta$, und dies bedeutet eben, daß (5_1) auch für $\mathfrak{T}$ gilt.

Um die Restbehauptung $|T|<|S|$ zu beweisen, schreiben wir (1) in der Form

$$\omega=\frac{\gamma}{\alpha}+\frac{\overline{\omega}}{\alpha(\alpha+\beta\overline{\omega})}=\frac{\delta}{\beta}-\frac{1}{\beta(\alpha+\beta\overline{\omega})}.$$

Wie eben erwähnt, ist $\operatorname{sgn}\overline{\omega}=\operatorname{sgn}\alpha\beta$, und so folgt, daß ω zwischen $\frac{\gamma}{\alpha}$ und $\frac{\delta}{\beta}$ liegt. Wegen (7) ist wenigstens das eine von $\left|r-\frac{\gamma}{\alpha}\right|$ und $\left|r-\frac{\delta}{\beta}\right|$ kleiner als 1. Wäre das andere größer als 1, so läge eine ganze rationale Zahl (nämlich $r+1$ oder $r-1$) zwischen

$\frac{\gamma}{\alpha}$ und $\frac{\delta}{\beta}$. Hieraus folgt, daß es zwischen $\beta\gamma$ und $\alpha\delta$ ebenfalls eine ganze rationale Zahl liegt, das doch wegen $\alpha\delta = \beta\gamma + 1$ unmöglich ist. Also ist immer

$$\left| r - \frac{\gamma}{\alpha} \right| \leqq 1, \quad \left| r - \frac{\delta}{\beta} \right| \leqq 1,$$

und es gilt hier wenigstens ein Zeichen $<$. Nach Multiplizieren mit $|\alpha|$ bzw. $|\beta|$ und Addieren folgt wegen (8)

$$(9) \qquad |\gamma'| + |\delta'| < |\alpha| + |\beta|.$$

Nach Obigem muß wegen $|\omega| > 1$ wenigstens das eine von $\left|\frac{\gamma}{\alpha}\right|$ und $\left|\frac{\delta}{\beta}\right|$ größer als 1 sein. Wäre das andere kleiner als 1, so würde wieder die Unmöglichkeit folgen, daß zwischen $\frac{\gamma}{\alpha}$ und $\frac{\delta}{\beta}$ eine ganze rationale Zahl liegt. Also ist

$$|\alpha| + |\beta| < |\gamma| + |\delta|.$$

Dies mit (8) und (9) ergibt $|T| < |S|$, womit der Satz bewiesen ist.

Unser Satz enthält den obigen Satz von GAUSZ offenbar. Wir zeigen noch, wie man aus unserem Satz (wie gesagt, ebenfalls ohne Kettenbrüche) zur Fundamentallösung der Pellschen Gleichung

$$(10) \qquad t^2 - D u^2 = 1$$

kommt. Der Spezialfall $\overline{\varphi} = \varphi$ ergibt nämlich, daß alle Substitutionen S, die eine reduzierte Form φ in sich überführen, durch

$$(11) \qquad S = \pm (E_1 E_2 \ldots E_m)^i \qquad (i = 0, \pm 1, \pm 2, \ldots)$$

angegeben sind. Diese und die Lösungen von (10) stehen miteinander in einem bekannten ein-eindeutigen Zusammenhang. Ein Teil davon ist

$$\gamma = a u.$$

Die Bedingung für nicht triviale Lösungen ($u \neq 0$) von (10) lautet also, daß es unter den S in (11) eins mit $\gamma \neq 0$ gibt. Nun ist $S_0 = E_1 E_2 \ldots E_m$ ein solches, was man so beweist. Wir setzen

$$E_k = \begin{pmatrix} 0 & 1 \\ -1 & \delta_k \end{pmatrix}.$$

Dann haben $\delta_1, \delta_2, \ldots$ abwechselnde Vorzeichen. Betrachten wir z. B. den Fall $\delta_1 > 0$ (der Fall $\delta_1 < 0$ ließe sich ähnlich behandeln).

Dann sind alle vier Elemente von

$$E_1 E_2 = \begin{pmatrix} -1 & \delta_2 \\ -\delta_1 & -1 + \delta_1 \delta_2 \end{pmatrix}$$

von gleichem (negativem) Vorzeichen. Ähnliches gilt für $E_3 E_4$, $E_5 E_6, \ldots$ und also, da m gerade ist, offenbar auch für S_0. Das beweist schon, daß S_0 eine nicht triviale Lösung von (10) liefert. Weiter folgt, daß die Absolutwerte aller vier Elemente in S_0^i mit i monoton zunehmen, und somit entspringt aus $S = S_0$ (vom Vorzeichen abgesehen) eben die Fundamentallösung von (10).

(Eingegangen am 26. November 1941.)

Rectification au travail "Sur la théorie ergodique des espaces abstraits"[1]).

Par F. RIESZ à Szeged.

Mon attention vient d'être attirée à ce que le N° **III. 3,** sous sa présente forme, est dépourvu de sens, dû à un oubli dans l'énoncé aussi bien qu'à un désarroi des détails. Voici comment il doit être corrigé.

Page 11, ligne 8 en remontant, au lieu de *et,* lire *si, et seulement si, $f^{(\nu)} \to f$, et alors on a.*

Page 11, ligne 3 en remontant, *au lieu de* D'autre part, *lire* Or, lorsque f est quarrable.

Page 12, ligne 6, *au lieu de* Enfin, *lire* Inversement, supposons que $f^{(\nu)} \to f$, alors.

Page 12, ligne 8, *au lieu de* ce qui précède, *lire* l'hypothèse faite.

Après cela, pour mieux comprendre le N° **4,** on fera bien d'intercaler page 12, ligne 7 en remontant, après la formule:

De plus, en raison de (2) et (4),

$$|f-f^{(\nu)}| \leqq |f-f_n| + |f_n - f_n^{(\nu)}| + |f_n^{(\nu)} - f^{(\nu)}| \leqq 2|f-f_n| + |f_n - f_n^{(\nu)}|;$$

comme $|f-f_n| \to 0$ d'après l'hypothèse et comme de plus, d'après le N° précédent, $|f_n - f_n^{(\nu)}| \to 0$ pour n fixé et $\nu \to \infty$, on a aussi $f^{(\nu)} \to f$.

(Reçu le 18 juin 1941)

---

1) *Ces Acta,* **10** (1941), p. 1–20.

Bibliographie.

E. Tornier, Wahrscheinlichkeitsrechnung und allgemeine Integrationstheorie, VI + 160 S, Leipzig und Berlin, B. G. Teubner, 1936.

Die durch v. Mises' Häufigkeits-Grenzwerttheorie 1919 heraufbeschworene Grundlagenkrise der Wahrscheinlichkeitsrechnung kann, wie Nachklänge der Genfer Konferenz 1937 und neuere Arbeiten zeigen noch immer nicht als überwunden betrachtet werden. Die Notwendigkeit einer zeitgemäßen Erneuerung der erschütterten klassischen Grundlagen wird zwar bereits allgemein anerkannt, in bezug auf Methode und Ausmaß des Neubaus ist aber noch starke Divergenz feststellbar. Auf die Methode würde es nicht so sehr ankommen, falls die *synthetische* Methode ein System der durch die *axiomatische* Methode zusammengefassten darstellen würde. Torniers Untersuchungen, die seit 1936 in Buchform erweitert vorliegen, lassen aber deutlich erkennen, daß dies z. B. bei der v. Mises-Waldschen synthetischen und der Kolmogoroffschen axiomatischen nicht der Fall sein kann. Soll nämlich die Wahrscheinlichkeit einerseits als Grenzwert der relativen Häufigkeit von Wiederholungserscheinungen, andererseits als Mengenfunktion über den Merkmalraum solcher Erscheinungen, gedeutet werden, so kann eben diese Funktion, wie die beim Peano-Jordanschen Inhalt herangezogene, nur *endlich additiv,* nicht aber, wie jene beim — von KOLMOGOROFF eingeführten — Lebesgueschen Maß, *totaladditiv* sein.

Der zu dieser Einsicht nötigen allgemeinen Theorie der Mengenfunktionen ist der erste, gut Zweidrittel des Ganzen und viele originelle Beiträge enthaltende Teil des Tornierschen Buches gewidmet, während der Rest einer im obigen Sinne doppelt deutbaren Wahrscheinlichkeitsrechnung des Verfassers vorbehalten bleibt. Letztere kann allerdings — wie die erwähnte Inhaltstheorie der allgemeinen Maßtheorie — nur als eine Spezialisation der häufigkeitstheoretisch zwar nicht deutbaren, aber wesentlich umfangreicheren und doch übersichtlicheren Kolmogoroffschen Theorie aufgefasst werden.

T. v. Stachó.

E. Kähler, Einführung in die Theorie der Systeme von Differentialgleichungen (Hamburger math. Einzelschriften, Heft 16), IV + 79 S., Leipzig und Berlin, B. G. Teubner, 1934.

Nachdem die Frage nach der Existenz der Lösungen von beliebig vielen Differentialgleichungen mit beliebig vielen Variabeln von RIQUIER und TRESSE auf dem von CAUCHY gewiesenen Wege erledigt worden ist, haben E. CARTAN und GOURSAT eine umfassendere Integrationstheorie von solchen Differentialgleichungssystemen geschaffen, welche durch Annullieren von alternierenden Differentialformen beliebigen Grades entstehen.

(Die Differentialformen ersten Grades nennt man bekanntlich auch Pfaffsche Formen.) Verfasser will eine systematische Einführung in die Theorie der Systeme von Differentialgleichungen geben, wie sie sich bei konsequenter Verwendung der alternierenden Differentialformen gestaltet.

Es werden zunächst die Addition, Multiplikation und Differentiation der alternierenden Differentialformen erklärt und die Koordinateninvarianz des Kalküls bewiesen. Im nächsten Kapitel folgen funktionentheoretische und geometrische Zwischenbetrachtungen. Die Hauptergebnisse der Theorie werden in zwei Existenzsätzen ausgesprochen, welche im wesentlichen aussagen, daß eine mehrdimensionale Integralmannigfaltigkeit sukzessiv aus niedrigerdimensionalen aufgebaut werden kann.

Als Anwendung werden die vollständig integrierbaren Pfaffschen Systeme und die Zurückführung partieller Differentialgleichungen auf skalare und Pfaffsche Systeme behandelt. Im Anhange sind die Hauptsätze der Lieschen Gruppentheorie zusammengestellt.

E. Egerváry.

Rudolf Weyrich, Die Zylinderfunktionen und ihre Anwendungen, VI + 137 S., Leipzig und Berlin, B. G. Teubner, 1937.

Das vorzügliche Büchlein enthält alles in der Theorie und Anwendung landläufig Gebrauchte. Von den — für spezielle Fälle aus Partikularlösungen der Wellenformgleichung durch transzendente Überlagerung erhaltenen — Sommerfeldschen Integraldarstellungen als Definitionen ausgehend, durch die Potenz- bzw. (sowohl Hankelschen als Debyeschen) asymptotischen Reihen, die im Zeiger bzw. in der Veränderlichen algebraischen und transzendenten Funktionalgleichungen, sowie die Diskussion des Funktionenverlaufs, um erst bei den Reihen- und Integraldarstellungen willkürlicher Funktionen durch Zylinderfunktionen referierend, mit Beispielen aus der Mechanik, sowie Feld- und Wärmetheorie zu schließen.

Bildliche Darstellungen fehlen zwar mit Berufung auf das — im Text erläuterte — Jahnke-Emdesche Tafelwerk, die auf Elemente der reellen und komplexen Funktionentheorie beschränkte mathematische Darstellung ist aber anschaulich und korrekt, die Angabe der Literatur und die Besprechung der dort zu findenden mannigfachen Bezeichnungen mustergültig.

T. v. Stachó.

Joseph Miller Thomas, Differential Systems (American Math. Society Colloquium Publications, Volume XXI), IX + 118 pages, New York, American Mathematical Society, 1937.

This valuable work deals with the theory of systems of partial differential equations as developed by CARTAN. This theory concerns the existence of solutions. Many personal contributions of the author are included. The treatment is *axiomatic*, sometimes even at the expense of perspicuity. The author aims at the highest degree of accuracy. Conciseness renders possible to include a great quantity of matter.

A concise summary of Grassmann algebra and elimination theory of algebraic equations is given in the introductory chapters. Beyond the axiomatic treatment, the work differs essentially from others dealing with the same subject by basing the theory upon systems of algebraic equations and "inequations." The exhaustive treatment of algebraic differential systems is a model, from which a short chapter leads the reader to general differential systems. Examination of the integral varieties of non-linear pfaffian systems (introduced by GOURSAT) completes the enormous matter.

At the end consistency examples are given, i. e. proofs of the existence theorems, on which as assumptions the axiomatic treatment has been based. Finally illustrative examples throw light upon some details. Reviewer thinks that a few more of these would have greatly increased the clearness.

G. Hajós.

Gustav Doetsch, Theorie und Anwendung der Laplace-Transformation (Grundlehren der math. Wissenschaften, Band XLVII), XVI + 436 S, Berlin, J. Springer, 1937.

Es liegt uns hier *die erste und zugleich eine vorzügliche Monographie* des — über die positive Achse erstreckten — Integrals $\int a(t)e^{-st}dt = f(s)$, und zwar bei nach RIEMANN absolut integrablen $a(t)$ als Funktionentransformation $L(a(t)) = f(s)$ aufgefasst vor. Daß die Monographie dieses, bereits von EULER und insbesondere von LAPLACE zur Lösung von Differential- bzw. Differenzengleichungen angewandten Integrals, das recht frühzeitig als das stetige Analogon der Potenzreihe $\Sigma a_n e^{-sn}$ erkannt und in den letzten Jahrzehnten durch die und neben den Dirichletschen Reihen $\Sigma a_n e^{-s\lambda_n}$ gefördert wurde, so lange auf sich warten ließ, erklärt sich daraus, daß es weiten und scharfen Blickes, sowie besonderen Geschickes bedürfte, das über dasselbe zerstreute Material zu sammeln, und richtiges vom falschen trennend gar zu oft oberflächlichen Analogieschlüssen oder formalen Rechnungen eine sichere Grundlage zu schaffen.

Herr G. DOETSCH, der die Laplace-Transformation zur Lösung von Integro- bzw. partiellen Differentialgleichungen, zur Erforschung von tiefliegenden Funktionengleichungen, zur Untersuchung von divergenten Reihen und asymptotischen Entwicklungen mit zeitgemäßer Strenge in 30 Abhandlungen so erfolgreich heranzog, löste diese schwere, aber dankbare Aufgabe vollkommen. Sein Werk stellt in systematischem Aufbau alles dar, was sich um seine Untersuchungen gruppieren läßt und zwar mit eingehenden geschichtlichen Anmerkungen, sowie einem reichen Literaturverzeichnis versehen.

Der erste Teil bringt die Grundlagen, d. h. Sätze über Gegenstand- bzw. Bildbereich der Transformation, ihre Umkehrung und die durch sie bewirkte Abbildung von fundamentalen Operationen. Der zweite, etwas kürzere, die Anwendung auf die Ableitung von konvergenten Reihenentwicklungen und eine sehr beachtenswerte systematische Behandlung bisher zerstreuter Betrachtungen über asymptotisches Verhalten von Funktionen. Das letzte Drittel ist den reichen Anwendungen auf die Lösung (bzw. Feststellung) von Funktionengleichungen gewidmet. Insbesondere werden hier Integral- und Integrodifferentialgleichungen vom Faltungstyp, sowie gewöhnliche und partielle Differentialgleichungen mit konstanten Koeffizienten besprochen und zwar am Beispiel der Wärme-, der Telegraphen- und der Potentialgleichung auch die feinsten Einzelheiten der Anfangs- und Randwertprobleme beleuchtend. Ausblicke auf Gleichungen mit veränderlichen Koeffizienten, auf die symbolische und funktionentheoretische Methode der Techniker, sowie eine Tabelle von Transformationsformeln beschließen dann das überaus reichhaltige Werk.

Als Kenner und Forscher des Gebietes zollen wir dem viele originelle Beiträge enthaltenden Werk volle Anerkennung und sehen seiner versprochenen, technischen Anwendungen bestimmten Fortsetzung gespannt entgegen.

T. v. Stachó.

Siegfried Valentiner, Vektoranalysis (Sammlung Göschen, 354), fünfte, erneut durchgesehene Auflage, 136 S., Berlin, Walter de Gruyter, 1938.

Die vorliegende fünfte Auflage des Büchleins enthält nur sehr wenig Änderungen gegenüber der letzten. Die verhältnismäßig ungewöhnt große Anzahl der Auflagen zeigt, daß die geschickt zusammengestellte kurze Darstellung der Vektoranalysis und einiger ihrer Anwendungen auf physikalische Probleme die verdiente Beachtung gefunden hat.

Béla v. Sz.-Nagy.

D. Hilbert und W. Ackermann, Grundzüge der theoretischen Logik (Grundlehren der math. Wissenschaften, Band XXVII), zweite Auflage, VIII + 133 S., Berlin, J. Springer, 1938.

Auf dem Titelblatt heißt es: zweite *verbesserte* Auflage; jedoch war im vorzüglichen und beliebten Werk — abgesehen von einer ungenauen Formulierung der Schlußregeln für den engeren Funktionenkalkül in der ersten Auflage — wohl nichts zu verbessern. Jedenfalls wurden die Fachausdrücke modernisiert und zugleich dem Hilbert-Bernaysschen Werk: *Grundlagen der Mathematik* angepaßt. Eine Unterscheidung zwischen freien und gebundenen Variablen in der Bezeichnung wurde im Gegensatz mit jenem Werk nicht unternommen.

Weggelassen wurde der Abriß der verzweigten Typentheorie, die ja nicht mehr diejenige Beachtung erfährt, die ihr zur Zeit der Erscheinung der ersten Auflage noch allgemein gewährt wurde. Dafür wurden die neueren Ergebnisse in der axiomatischen Erforschung des Aussagen- und

Prädikatenkalküls, sowie in der Reduktions- und Lösungstheorie des Entscheidungsproblems, auch betreffend die Unlösbarkeit desselben, *berücksichtigt.*

Das Buch ist auch heute in keiner Hinsicht von einem Werk gleichen Umfangs übertroffen.

L. Kalmár.

Gerhard Kowalewski, Grundbegriffe und Hauptsätze der höheren Mathematik insbesondere für Ingenieure und Naturforscher, 156 S., Berlin, Walter de Gruyter, 1938.

Wer es bereits versucht hat, Differential- und Integralrechnung denjenigen verständlich zu machen, die sie nicht aus rein wissenschaftlichem Interesse, sondern in erster Linie um der Anwendungen willen studieren, weiß aus Erfahrung, mit welchen Schwierigkeiten diese Aufgabe behaftet ist. Verzichtet man auf die Exaktheit, so besteht die Gefahr, daß der Student sein Vertrauen in den mathematischen Methoden verliert, was keineswegs das Verständnis ihrer Anwendungen fördert; beharrt man dagegen darauf, so fällt man leicht in die Versuchung, in schwierige Einzelheiten einzugehen, die wohl für den systematischen Aufbau, nicht aber für die Anwendungen wichtig sind.

Die lange Erfahrung Kowalewskis einerseits als Lehrer, andererseits als Verfasser von Lehrbüchern hat ihm auch bei der Lösung dieser schwierigen Aufgabe als Wegweiser gedient. Ohne ein Haar breit in Strenge nachzugeben, beschränkt er sich in denjenigen Teilen der Theorie, die für die Anwendungen belanglos sind, auf das notwendigste. Nach einer Vorbereitung über Koordinaten, Vektoren und Determinanten entwickelt er die Lehre von den Grenzwerten und die für die Anwendungen unentbehrlichen Grundkenntnisse aus der eigentlichen Differential- und Integralrechnung. Die Anschauung wird nicht nur zur Erleichterung schwieriger Gedankengänge herangezogen, sondern die ganze Darstellung, der Stil des Verfassers ist von Anschaulichkeit durchdrungen. Oft werden längere Ausführungen durch einen gutgewählten, anschaulichen Fachausdruck erspart.

Einige Übungsbeispiele, insbesondere aus den Anwendungsgebieten, würden den Wert des sonst vorzüglichen, auch für Studierende der Mathematik als erste Einführung sehr brauchbaren Lehrbuches wesentlich erhöhen.

L. Kalmár.

Gabor Szegő, Orthogonal Polynomials (American Math. Society Colloquium Publications, Volume XXIII), IX + 401 pages, New York, American Mathematical Society, 1939.

Orthogonal polynomials are a pet subject of the author, right from the beginning. While reading his present book, my mind wanders back to 1924, when, acting as a J. König's Prize referee, I had to report on his work, the first few years of his work indeed. I remember now, how I was struck by the simplicity and originality, in particular, of two of his papers, in which, starting from a most elementary theorem of Fejér's on the normalized representation of positive trigonometric polynomials, he developed a much simple but powerful method to deal with the asymptotic behaviour of orthogonal polynomials, of some very general types as well as with that of the "kernels" occuring in the corresponding expansion problems. Let me also mention a paper of his to which I added some remarks of my own; it dealt with functions analytical in the unit circle, with given modulus on the circle itself, and its results belong to the present subject as an important tool.

These were the beginnings of a life-work that led up to the present standard treatise on orthogonal polynomials. that is, I forgot to say, polynomials orthogonalized, with respect to a given mass distribution, on a real interval, finite or infinite, as well as on more general curves, on the unit circle in particular. The best known instances for the first case are the Jacobi polynomials, including the Legendre ones, further those of Hermite and Laguerre. The case of the unit circle includes, as its simplest item, the set (z^n). n running from $-\infty$ to $+\infty$ and with it, in an obvious sense, the classical Fourier set too. Orthogonal polynomials are related to hypergeometric, Bessel, and elliptic functions, to interpolation and mechanical quadrature. The present work is devoted partly to the study of special classes and to their applications, partly to the general theory, the author keeping, as far as possible, to elementary methods and general principles and entering upon more hardworking computation only when need arises, but even than taking care not to hide the essential ideas.

I am glad to recommend to the mathematical public the present standard work of a mathematician, teaching abroad, but born and educated in this country.

F. R.

Constantin Carathéodory, Reelle Funktionen, Band I: Zahlen, Punktmengen, Funktionen, VI + 184 S., Leipzig und Berlin, B. G. Teubner, 1939.

Das vorliegende Buch ist der erste Teil einer wesentlich neuen Bearbeitung der alten *Vorlesungen über reelle Funktionen* aus 1918. Der Inhalt entspricht ungefähr den ersten 229 Seiten des alten Buches: reelle Zahlen, Grenzbegriff, das wesentlichste über Punktmengen, stetige, halbstetige, punktiert und totalstetige und monotone Funktionen, stetige Abbildung und Funktionenfolgen. Wesentlich umgestaltet ist die Anordnung der Sätze und von den sich darbietenden Beweismethoden werden die erweiterungsfähigen bevorzugt, um so spätere Verallgemeinerungen vorzubereiten. Neu hinzugekommen sind ein integralloser, auch für allgemeinere topologische Räume gültiger, im wesentlichen auf TIETZE, 1915, zurückgehender Beweis von R. RADO für die Erweiterung des Definitionsbereichs von stetigen

Funktionen, ferner die der gleichmäßigen nahe verwandte „stetige“ Konvergenz und die normalen Familien von Funktionen.

Mit gespannter Erwartung sehen wir dem zweiten Bande entgegen, der die im alten Werke meisterhaft aufgebaute Maß- und Integrationstheorie enthalten soll, voraussichtlich in einer durch die allerneuesten Untersuchungen des Verfassers und Anderer über die Algebraisierung dieser Begriffe beeinflußten Darstellung.

F. R.

B. L. van der Waerden, Moderne Algebra, zweiter Teil (Grundlehren der math. Wissenschaften, Band XXXIV), zweite verbesserte Auflage, VIII + 224 S., Berlin, J. Springer, 1940.

Die vorliegende zweite Auflage des zweiten Bandes des ausgezeichneten Werkes zeigt gegenüber der ersten folgende Änderungen auf: Die Paragraphen über Resultanten, da sie in der zweiten Auflage ins erste Band übernommen wurden, fallen jetzt natürlich weg. Dafür wurden die beiden letzten Kapitel, die hyperkomplexe Größen und ihre Darstellungen behandeln, entsprechend der stürmischen Entwicklung der neueren Algebrentheorie, erheblich erweitert und umgestaltet. Aber auch die übrigen Kapitel wurden einer sorgfältigen Revision unterzogen. Alles in Allem ist die zweite Auflage der *Modernen Algebra* jetzt ebenso modern, wie die erste Auflage vor zehn Jahren war, die Darstellung ist aber noch vollständiger geworden.

Béla v. Sz. Nagy.

Zur Begründung der Minkowskischen Geometrie.

Von O. VARGA in Kolozsvár.

Obgleich die Minkowskische Geometrie ein Spezialfall der in den letzten Jahren von verschiedenen Autoren behandelten Finslerschen Geometrie ist, scheint es doch wünschenswert eine direkte Herleitung dieser Geometrie zu geben. Dies umso eher, da die Minkowskische Geometrie in viel näherer Beziehung zur Euklidischen Geometrie steht, als dies von der Finslerschen gilt.

Der grundlegende Gedanke, der in dieser Arbeit verwendet wird, ist der folgende: der Minkowskische Raum ist ein affiner Raum, in dem jeder Richtung eine euklidische Maßbestimmung zugeordnet ist. Dementsprechend behandeln wir §. 1 den Minkowskischen Raum in Bezug auf ein kartesisches Koordinatensystem. Im zweiten Paragraphen führen wir krummlinige Koordinaten ein. Es ist bemerkenswert, daß man dann den ganzen formalen Apparat der Finslerschen Geometrie erhält. Dies ist ein Analogon zur Tatsache, daß der auf krummlinige Koordinaten bezogene Euklidische Raum den formalen Apparat der Riemannschen Geometrie liefert. Wie E. Cartans Behandlungsweise der Riemannschen Geometrie zeigt[1], sind diese Zusammenhänge nicht bloß formal bedingt, sondern haben einen weitgehenden inneren Grund. Derselbe ermöglicht E. CARTAN die einfache und elegante Herleitung der Riemannschen Geometrie[2]. Der entsprechende Aufbau der Finslerschen Geometrie mit Hilfe der Minkowskischen ist ebenfalls möglich und soll an anderer Stelle behandelt werden[3].

---

[1]) E. CARTAN (1), S. 39—59. (S. Schriftenverzeichnis am Ende vorliegender Arbeit.)

[2]) E. CARTAN (1), S. 90—180.

[3]) O. VARGA (2).

Die Krümmungstheorie wird in §. 3 nur ganz kurz berührt, da bei ihrer Behandlung fast kein Unterschied gegenüber den Gedankengängen zu verzeichnen ist, die zu denselben Zwecken in der Finslerschen Geometrie verwendet werden. Als Anwendung der Krümmungstheorie weisen wir nach, daß der Minkowskische Raum unter den Finslerschen durch das Verschwinden zweier Tensoren ausgezeichnet ist. Diese Charakterisierung wurde ohne Beweis von E. CARTAN angegeben[4]).

§. 1. Der Minkowskische Raum in Bezug auf ein kartesisches Koordinatensystem.

1. Definition des Minkowskischen Raumes und Zusammenhang mit dem Euklidischen Raum. Eine n-dimensionale Mannigfaltigkeit heißt ein Minkowskischer Raum, wenn es in ihr ein System von Koordinaten $x^1, x^2, \ldots, x^n$ gibt, so daß die Entfernung zweier benachbarter Punkte (x) und $(x+dx)$ durch

$$ds = L(dx^1, dx^2, \ldots, dx^n) \tag{1, 1}$$

oder kürzer

$$ds = L(dx) \tag{1, 1'}$$

gegeben ist. Von $L(dx)$ soll vorausgesetzt werden, daß es beständig positiv, ferner in den $dx^1, dx^2, \ldots, dx^n$ positiv-homogen von erster Dimension ist und nach diesen Veränderlichen stetige Ableitungen bis zur 4. Ordnung besitzt. Die quadratische Form[5])

$$\frac{\partial^2 \left(\frac{1}{2} L^2(dx)\right)}{\partial dx^i \partial dx^k} Z^i Z^k \tag{1, 2}$$

der Hilfsveränderlichen $Z^i (i = 1, 2, \ldots, n)$ sei für alle Wertesysteme der $dx^1, dx^2, \ldots, dx^n$ positiv definit.

Geht man von dem Koordinatensystem $(x^1, x^2, \ldots, x^n)$ durch eine beliebige stetig differenzierbare und umkehrbare Transformation

$$x^i = x^i(u^1, u^2, \ldots, u^n) \qquad (i = 1, 2, \ldots, n) \tag{1, 3}$$

zu einer Koordinatenbestimmung $(u^1, u^2, \ldots, u^n)$ über, so wird in derselben das Bogenelement wegen (1, 1) die Form

---

[4]) E. CARTAN (2), S. 39.

[5]) Tritt ein Zeiger in einem Ausdruck mehr als einmal auf, so soll nach demselben — ohne Angabe eines Summationszeichens — summiert werden.

$$(1, 4) \qquad ds = L\left(\frac{\partial x^1}{\partial u^s} du^s, \frac{\partial x^2}{\partial u^s} du^s, \ldots, \frac{\partial x^n}{\partial u^s} du^s\right) \equiv$$
$$\equiv \mathsf{L}(u^1, u^2, \ldots, u^n, du^1, du^2, \ldots, du^n)$$

annehmen. In dem neuen Koordinatensystem wird also das Bogenelement insofern durch einen komplizierteren Ausdruck angegeben, da außer den Differentialen (du) auch die Punktkoordinaten (u) auftreten.

Wir wollen ein Koordinatensystem, in dem das Bogenelement nur von den Differentialen abhängt, als ein ausgezeichnetes bezeichnen. Aus (1, 3) und (1, 4) folgt sofort, daß man aus einem ausgezeichneten Koordinatensystem sämtliche übrige durch lineare Transformationen erhält. Im Falle einer Translation bleibt die Funktion $L(dx)$ ungeändert.

Wir wollen einen Minkowskischen Raum — bei Verwendung eines ausgezeichneten Koordinatensystems — als einen *affinen Raum* auffassen, dem durch (1, 1) eine Maßbestimmung aufgeprägt ist. Das verwendete Koordinatensystem soll daher statt ausgezeichnet als *kartesisch* bezeichnet werden. Eine für späteres wichtige Feststellung ist die folgende: Ist ein Vektorfeld gegeben, so ist die Ableitung des Feldvektors in allen kartesischen Koordinatensystemen wieder ein Vektor. Allgemeiner wird der tensorielle Charakter von Größen in kartesischen Koordinatensystemen durch Differentiation nicht gestört. Dies folgt unmittelbar aus der Definition eines Tensors durch Transformationsgleichungen und aus dem Umstand, daß man von einem kartesischen Koordinatensystem zu einem anderen durch lineare Transformationen gelangt.

Ausgehend von (1, 1) kann man auch den Abstand $d(x_0 x_1)$ eines beliebigen Punktes (x_1) von (x_0) erklären. Man betrachtet dazu die orientierte Gerade, die von x_0 nach x_1 läuft. Ist

$$(1, 5) \qquad x^i = x_0^i + a^i(t - t_0) \qquad (a^i > 0,\ i = 1, 2)$$

ihre Parameterdarstellung und entspricht dem Punkte x_ϱ der Parameterwert t_ϱ $(\varrho = 0, 1)$, so soll $d(x_0, x_1)$ durch

$$(1, 6) \qquad d(x_0, x_1) = \int_{t_0}^{t_1} L\left(\frac{dx}{dt}\right) dt$$

definiert sein. Wegen (1, 5) und der Homogenität von $L(dx)$ ergibt sich aus (1, 6)

(1, 7) $$d(x_0, x_1) = L(x_1 - x_0).$$

Der Vektor ξ^i mit dem Komponenten $x^i - x_0^i$ ist daher dann und nur dann ein Einheitsvektor, wenn

(1, 8) $$L(\xi) = 1$$

besteht.

Trägt man von einem beliebigen Punkt (x_0) nach allen Richtungen Einheitsvektoren ab, so ist der Ort ihrer Endpunkte (x) die Hyperfläche

(1, 9) $$L(x - x_0) = 1$$

die Indikatrix oder Eichfläche des Punktes (x_0). Die Indikatrizen verschiedener Punkte gehen durch Translation auseinander hervor. Wegen (1, 2) ist die Indikatrix eine konvexe Hyperfläche.

Wir wollen jetzt jeder Richtung des Minkowskischen Raumes eine Euklidische Maßbestimmung zuordnen. Dazu betrachten wir eine beliebige, aber feste Richtung $\dot{x}_0^i$ durch einen ebenfalls willkürlichen Punkt x_0^i. Die Komponenten des Einheitsvektors dieser Richtung sind wegen (1, 8)

(1, 10) $$\xi_0^i = \frac{\dot{x}_0^i}{L(\dot{x}_0)} \qquad (i = 1, 2, \ldots, n).$$

Es gibt nur eine einzige Mittelpunktsfläche zweiter Ordnung, die den Punkt x_0^i zum Mittelpunkt hat und die Indikatrix des Punktes x_0^i im Punkte mit dem Koordinaten

(1, 11) $$x^i = x_0^i + \xi_0^i \qquad (i = 1, 2, \ldots, n)$$

oskuliert. Sie hat in laufenden Koordinaten $X^1, X^2, \ldots, X^n$ die Gleichung

(1, 12) $$g_{ik}(\dot{x}_0)(X^i - x_0^i)(X^k - x_0^k) = 1,$$

wobei

(1, 13) $$g_{ik}(\dot{x}) = \frac{\partial^2 \left(\frac{1}{2} L^2(\dot{x})\right)}{\partial \dot{x}^i \partial \dot{x}^k}$$

ist. Die $g_{ik}(\dot{x})$ sind in beiden Zeigern symmetrisch und von nullter Dimension positiv-homogen in den $\dot{x}^i$. Sie hängen also nur von der Richtung des Vektors $\dot{x}^i$ und nicht von seiner Länge ab.

Die Hyperfläche (1, 12) heißt die oskulierende Indikatrix oder oskulierende Eichfläche des Punktes (x_0) nach der Richtung $\dot{x}_0^i$. Die oskulierenden Indikatrizen verschiedener Punkte nach der gleichen Richtung $\dot{x}_0^i$ gehen durch Translation auseinander hervor. Bei festgehaltener Richtung $\dot{x}_0^i$ ist die Maßbestimmung

$$(1,14) \qquad d\sigma^2 = g_{ik}(\dot{x})\, dx^i\, dx^k,$$

deren Indikatrix in jedem Punkte x_0 die dort nach Richtung $\dot{x}_0^i$ oskulierende Indikatrix ist, eine Euklidische Maßbestimmung, die der Richtung $\dot{x}_0^i$ zugeordnet ist. Sie heißt die nach der Richtung $\dot{x}_0^i$ oskulierende Maßbestimmung des Minkowskischen Raumes. Sie ergibt natürlich nur für Vektoren der Richtung $\dot{x}_0^i$ dieselbe Länge wie die Maßbestimmung (1, 1).

Durch die oskulierende Maßbestimmung wird jeder Richtung $\dot{x}^i$ des Minkowskischen Raumes ein Euklidischer Raum zugeordnet. Er möge als der Euklidische Raum $(\dot{x})$ bezeichnet werden. Der oskulierenden Maßbestimmung entsprechen im Euklidischen Raum $(\dot{x})$ diejenigen kartesischen Koordinatensysteme, deren Maßvektoren[6] $\underset{(i)}{e_k}$ den Gleichungen

$$(1,15) \qquad \underset{(i)}{e_s}\,\underset{(k)}{e_s} = g_{ik}(\dot{x})$$

genügen. Aus diesen $\infty^{\frac{n(n-1)}{2}}$, der Richtung $\dot{x}^i$ zugeordneten kartesischen Koordinatensystemen greifen wir ein beliebiges heraus und *setzen fest, daß sich dasselbe in stetig differenzierbarer Weise mit der Richtung $\dot{x}^i$ ändern soll,* d. h. die Komponenten der Maßvektoren seien stetig differenzierbare Funktionen der $\dot{x}^1, \dot{x}^2, \ldots, \dot{x}^n$.

Es ist jetzt möglich, die Länge von Vektoren *in Bezug auf eine Richtung* $(\dot{x})$ zu erklären. Man versteht darunter die Länge l des Vektors $\vec{\xi}$ im zugeordneten Euklidischen Raum $(\dot{x})$

$$(1,16) \qquad l = \sqrt{g_{ik}(\dot{x})\,\xi^i \xi^k}.$$

Die Länge l ist unabhängig davon, in welchem Punkte der Vektor angebracht wird. Der Cosinus des Winkels ϑ zweier Vektoren $\vec{\xi}$ und $\vec{\eta}$ in Bezug auf die Richtung ist entsprechend gegeben durch

$$(1,17) \qquad \cos\vartheta = \frac{g_{ik}(\dot{x})\,\xi^i\eta^k}{\sqrt{g_{ik}(\dot{x})\,\xi^i\xi^k}\,\sqrt{g_{ik}(\dot{x})\,\eta^i\eta^k}}.$$

Allgemein wollen wir als skalares Produkt zweier Vektoren den folgenden Ausdruck bezeichnen:

$$(1,17') \qquad \vec{\xi}\,.\,\vec{\eta} = g_{ik}\,\xi^i\eta^k.$$

---

[6]) In der Bezeichnung $\underset{(i)}{e_k}$ bedeutet der in Klammer stehende Zeiger i den Vektor, der zweite Zeiger k gibt die Komponente an.

2. **Das invariante Differential.** Durch die oben getroffenen Festsetzungen ist ein Vektor wesentlich an die Richtung $(\dot{x})$ gebunden. Wir wollen nun festsetzen, daß wir unter Parallelverschiebung einer Richtung $\dot{x}^i$ die gewöhnliche Paralellverschiebung im affinen Raum $x^1, x^2, \ldots, x^n$ verstehen. Zwei parallelen Richtungen entsprechen somit proportionale $\dot{x}^i$. Entsprechend wollen wir unter der Parallelverschiebung eines Vektors $\vec{\xi}$ von einem Punkte des Raumes nach einem anderen, bei paralleler Richtung $(\dot{x})$, die gewöhnliche Parellelverschiebung im Euklidischen Raum $(\dot{x})$ verstehen. Wir wollen nun Festsetzungen treffen zum Vergleich zweier Vektoren, die zu verschiedenen Richtungen gehören. Diese Festsetzungen sollen infinitesimalgeometrisch gemacht werden. Durch den Feldvektor $\vec{V}(x, \dot{x})$ sei ein stetig differenzierbares Vektorfeld festgelegt. Dabei sei zugelassen, daß in $\vec{V}(x, \dot{x})$ die Variablen $x^1, x^2, \ldots, x^n$ nicht vorkommen, d. h., daß sämtliche Vektoren denselben Angriffspunkt haben. Sind $V^i(x, \dot{x})$ die Komponenten des Feldvektors, dann können wir den Vektor der Richtung $(x, \dot{x})$ sofort mit dem Vektor der willkürlichen benachbarten Richtung $(x+dx, \dot{x}+d\dot{x})$ vergleichen. Der Unterschied beider Vektoren wird durch $dV^i(x, \dot{x})$ angegeben, dies ist nämlich nach einer früher gemachten Bemerkung wieder ein Vektor. Nicht so einfach, für unsere Zwecke aber wichtiger gestaltet sich der Vergleich benachbarter Vektoren in den Komponenten $v^i(x, \dot{x})$ des Feldvektors in Bezug auf das lokale Koordinatensystem $\underset{(i)}{e_k}$, das der Richtung $(\dot{x})$ zugeordnet ist. Der Zusammenhang der v^i und V^i ist durch

(1, 18) $$V^i = \underset{(s)}{e_i} v^s$$

bestimmt. Unsere Aufgabe besteht nun darin, diejenigen Größen Dv^i zu bestimmen, die der Gleichung

(1, 19) $$dV^i = \underset{(s)}{e_i} Dv^s$$

genüge leisten. Differentiation der Gleichung (1, 18) zeigt, daß wir dazu nur die Größen $d\underset{(s)}{e_i}$ bestimmen, d. h. aus der Maßfunktion $L(dx)$ und ihren Ableitungen heraus berechnen müssen.

Nach der schon zuvor benützten Bemerkung sind die Größen $d\underset{(s)}{e_i}$ und ebenso natürlich $\dfrac{\partial \underset{(s)}{e_i}}{\partial \dot{x}^k}$ (bei festen s und k) Vektoren, daher lassen sie sich aus den $\underset{(\iota)}{e_i}$ linear kombinieren. Es ist also

$$(1,20)\qquad \frac{\partial \underset{(\iota)}{e_i}}{\partial \dot{x}^k} = C^r_{\iota k} \underset{(r)}{e_i}$$

mit Koeffizienten $C^r_{\iota k}$, die nur Funktionen von $\dot{x}^i$ sind, da wir ja den fraglichen Vektor im Ursprung des Koordinatensystems der $\underset{(r)}{e_i}$ — das nur von $\dot{x}^i$ abhängt — ansetzen. Aus (1, 15) ergibt sich ferner durch Differentiation und nochmaliger Benützung dieser Gleichungen selbst

$$(1,21)\qquad C^l_{ir} g_{lk} + C^l_{kr} g_{li} = \frac{\partial g_{ik}}{\partial \dot{x}^r}.$$

Um die Größen C^l_{ir} vollständig durch die Funktion $L(dx)$ und ihren Ableitungen zu bestimmen, stellen wir noch die folgende Symmetrieforderung

$$(1,22)\qquad C_{ikr} = C_{kir};$$

dabei haben wir

$$(1,23)\qquad C_{ikr} = C^l_{ir} g_{lk}$$

gesetzt. Jetzt ergibt sich aus (1, 21) und (1, 13)

$$(1,24)\qquad \boxed{C_{ikr} = \frac{1}{2}\frac{\partial g_{ik}(\dot{x})}{\partial \dot{x}^r} = \frac{1}{4}\frac{\partial^3 (L^2(\dot{x}))}{\partial \dot{x}^i \partial \dot{x}^k \partial \dot{x}^r}.}$$

Für $d\underset{(\iota)}{e_i}$ ergiebt sich somit

$$(1,25)\qquad d\underset{(\iota)}{e_i} = C^r_{\iota k} \underset{(r)}{e_i}\, d\dot{x}^k.$$

Differentiation von (1, 18) gibt somit bei Beachtung von (1, 25)

$$(1,26)\qquad dV^i = \underset{(s)}{e_i}(dv^s + C^s_{kl} v^k d\dot{x}^l).$$

Vergleich von (1, 19) und (1, 25) ergibt

$$(1,27)\qquad Dv^s = dv^s + C^s_{kl}\, v^k d\dot{x}^l.$$

Man bezeichnet Dv^s als das invariante Differential der Größen v^s.

Die Größen C_{ikr} sind, wie aus (1, 24) folgt, in allen drei Zeigern symmetrisch und in den $\dot{x}^i$ von (-1)-ter Ordnung positiv-homogen. Aus der letztgenannten Eigenschaft folgt noch

$$(1,28)\qquad C_{ikr}(\dot{x})\dot{x}^i = C_{ikr}\dot{x}^k = C_{ikr}\dot{x}^r = 0.$$

Da, wie oben bemerkt, Differentiation den tensoriellen Charakter einer Größe nicht ändert, sind die C_{ikr} ein Tensor dritter Stufe.

Bemerkung. Daß die Größen $\underset{(r)}{de_i}$ durch die $dg_{ik}(\dot{x})$ im wesetlichen bestimmt sind, ist wegen der Festlegung der Koordinatensysteme $\underset{(s)}{e_i}$ durch die g_{ik} ziemlich naheliegend. Die $\underset{(s)}{e_i}$ — genauer bloß ihre gegenseitige Lage — sind durch die g_{ik} bestimmt, die Nachbarvektoren $\underset{(s)}{e_i}+\underset{(s)}{de_i}$ daher durch den Nachbarwert der g_{ik}, also durch $g_{ik}+dg_{ik}$.

§. 2. Der Minkowskische Raum, auf krummlinige Koordinaten bezogen.

1. Einführung der krummlinigen Koordinaten und lokale Minkowskische Räume, die auf kartesische Koordinaten bezogen sind. In §. 1, 1 hatten wir, um die ausgezeichneten oder kartesischen Koordinaten definieren zu können, durch die Transformation (1, 3) Koordinaten $u^1, u^2, \ldots, u^n$ eingeführt, die wir als allgemeine oder auch krummlinige Koordinaten bezeichnen wollen. Wie wollen nun einen auf krummlinige Koordinaten bezogenen Minkowskischen Raum näher untersuchen. Zunächst einige Feststellungen über die Funktion $L(dx)$ und die aus ihr hergeleiteten Tensoren $g_{ik}(dx)$ und $C_{ikl}(dx)$ in den krummlinigen Koordinaten. Die Eigenschaft der Funktion $L(dx)$, in den Differentialen positiv-homogen von 1-ter Dimension zu sein, wird durch die Transformation (1, 3) nicht gestört, d. h. $\mathsf{L}(u, du)$ ist in den du^i positiv-homogen von 1-ter Dimension.

Die Größen $\dot{x}^i$, $g_{ik}(\dot{x})$, $C_{ikl}(\dot{x})$ sind Tensoren, transformieren sich also bei Ausübung von (1, 3) gemäß

$$(2,1)\qquad \begin{cases} \dot{x}^i = \dfrac{\partial x^i}{\partial u^k}\,\dot{u}^k, \\[2mm] g_{ik}(u, \dot{u}) = g_{rs}(\dot{x})\,\dfrac{\partial x^r}{\partial u^i}\,\dfrac{\partial x^s}{\partial u^k}, \\[2mm] C_{ikl}(u, \dot{u}) = C_{rsp}(\dot{x})\,\dfrac{\partial x^r}{\partial u^i}\,\dfrac{\partial x^s}{\partial u^k}\,\dfrac{\partial x^p}{\partial u^l}. \end{cases}$$

Die $g_{ik}(u, \dot{u})$ und $C_{ikl}(u, \dot{u})$ drücken sich durch die Funktion $\mathsf{L}(u, \dot{u})$ und ihre Ableitungen folgendermaßen aus:

$$(2,2)\qquad \begin{aligned} g_{ik}(u, \dot{u}) &= \frac{1}{2}\,\frac{\partial^2}{\partial \dot{u}^i \partial \dot{u}^k}\,(\mathsf{L}^2(u, \dot{u})), \\ C_{ikl}(u, \dot{u}) &= \frac{1}{4}\,\frac{\partial^3(\mathsf{L}^2(u, \dot{u}))}{\partial \dot{u}^i \partial \dot{u}^k \partial \dot{u}^l}. \end{aligned}$$

(2, 2) zeigt, daß die Größen $g_{ik}(u, \dot{u})$ und $C_{ikl}(u, u)$ in Bezug auf u^i von nullter bzw. (-1)-ter Dimension positiv-homogen sind. Die Symmetrieeigenschaft bleibt natürlich auch bestehen, da diese ja bei Tensoren koordinateninvariant ist. Ebenso besteht auch hier die Tensorrelation (1, 28).

Unser, auf ein kartesisches Koordinatensystem $x^1, x^2, \ldots, x^n$ bezogene Minkowskische Raum erscheint bei Ausübung der Transformation (1, 3) von einem Kurvennetz bedeckt, das natürlich durch die Linien

$$(2, 2) \qquad u^i = c \qquad (i = 1, 2, \ldots, n)$$

gebildet wird. In einem festen, aber beliebigen Punkt (u) des Raumes bestimmt dieses Kurvennetz durch seine Tangentenvektoren n Vektoren:

$$(2, 3) \qquad \underset{(i)}{e_s} = \frac{\partial x^s}{\partial u^i} \qquad (i = 1, 2, \ldots, n).$$

Andererseits wird durch Festhalten eines solchen beliebigen Punktes (u) demselben ein lokaler, auf kartesische Koordinaten bezogener Minkowskischer Raum zugeordnet, da ja dann in $L(u, \dot{u})$ nur die $\dot{u}^i$ variabel sind. Jedem Linienelement $(u, \dot{u})$ ist nun nach §. 1, 1 derjenige Euklidische Raum $(\dot{x})$ zugeordnet, dessen Komponenten $\dot{x}^i$ (2, 1) entsprechen. Wir wollen diesen Raum als den Euklidischen Raum $(u, \dot{u})$ bezeichnen. Als *Maßvektoren* dieses, auf ein kartesisches Koordinatensystem bezogenen Minkowskischen Raumes können wir die Vektoren (2, 3) verwenden. Dazu ist nur zu zeigen, daß das skalare Produkt zweier Vektoren $\underset{(i)}{e_i^\nu}$ (*im Koordinatensystem* $x^1, x^2, \ldots, x^n$!) die Größen $g_{ik}(u, \dot{u})$ ergeben. Der analytische Ausdruck hiefür wir aber bei Beachtung von (1, 17') und (2, 3) durch (2, 1) geliefert.

2. **Übertragung von Vektoren, das invariante Differential.** Nach §. 1, 2 bedeutet Parallelität zweier Richtungen Parallelität im Sinne der gewöhnlichen Euklidischen Geometrie. Wir betrachten nun in einem gewissen Bereich $\mathfrak{B}$ ein Feld von parallelen Richtungen. Die Gleichung des Feldvektors sei

$$(2, 4) \qquad u^i = \dot{u}^i(u^1, u^2, \ldots, u^n) \qquad (i = 1, 2, \ldots, n).$$

Die Funktionen (2, 4) sind dabei auf $\mathfrak{B}$ definiert. Die Überlegungen von § 1, 2 zeigen, daß in $\mathfrak{B}$ ein auf krummlinige Koordinaten $u^1, u^2, \ldots, u^n$ bezogener Euklidischer Raum vorliegt. Sein Maßtensor wird durch

$$(2,5)\quad \gamma_{ik}(u^1, u^2, \ldots, u^n) = \\ = g_{ik}(u^1, u^2, \ldots, u^n, \dot{u}^1(u^1, u^2, \ldots, u^n), \ldots, \dot{u}^n(u^1, u^2, \ldots, u^n))$$

bestimmt. In einem solchen Raum genügt aber der Feldvektor eines Parallelbündels (2, 4) der folgenden Differentialgleichung:

$$(2,6)\qquad d\dot{u}^i = \Gamma^{*\,i}_{kl}\dot{u}^k du^l.$$

Dabei sind die $\Gamma^{*\,i}_{kl}$ die mit Hilfe der γ_{ik} gebildeten Christoffelschen Symbole zweiter Art[7]).

Die Übertragung eines Vektors v^i von einem beliebigen Linienelement $(u, \dot{u})$ zu einem beliebigen benachbarten parallelen Linienelement $(u+du, \dot{u}+d\dot{u})$ — natürlich innerhalb eines Bereiches der Art $\mathfrak{B}$ — wird dann gleichfalls entsprechend dem in § 1, 2 Auseinandergesetzten und den Tatsachen, die in einem auf krummlinige Koordinaten bezogenen Euklidischen Raum bestehen, durch

$$(2,7)\qquad Dv^i = dv^i + \Gamma^{*\,i}_{kl} v^k du^l$$

ausgedrückt[7]). Der Parallelismus eines Vektorfeldes v^i bedeutet:

$$(2,8)\qquad Dv^i = 0.$$

Wir wollen jetzt die Übertragung eines Vektors von einem beliebigen Linienelement $(u, \dot{u})$ zu einem beliebigen benachbarten — nicht parallelen — Linienelement $(u+du, \dot{u}+d\dot{u})$ herleiten. Dies wird in zwei Schritten erfolgen. Erstens: Übertragung des Vektors v^i von $(u, \dot{u})$ zu dem benachbarten, aber parallelen Linienelement $(u+du, \dot{u}+d^*\dot{u})$. Zweitens: Übergang von dem Linienelement $(u+du, \dot{u}+d^*\dot{u})$ zu dem Linienelement $(u+du, \dot{u}+d\dot{u})$ gleichen Zentrums. Der erste Schritt vollzieht sich nach (2, 7), der zweite nach (1, 27). Wir wollen den analytischen Ausdruck, den wir durch diese Überlegungen gewinnen, herleiten. Bezeichnen wir das invariante Differential, das zu dem ersten Schritt gehört, mit $D_\pi v^i$ und mit $D_\varrho v^i$ das zum zweiten Schritt gehörige, so ergibt sich als Ausdruck für das invariante Differential

$$(2,9')\qquad Dv^i = D_\pi v^i + D_\varrho(v^i + D_\pi v^i),$$

oder, da wir Größen von höherer als erster Ordnung vernachläßigen,

$$(2,9)\qquad Dv^i = D_\pi v^i + D_\varrho v^i.$$

---

[7]) E. Cartan (1), S. 39, Gleichungen (15).

Nach (2, 7) gilt für $D_\pi v^i$:

(2, 10) $$D_\pi v^i = d^* v^i + \Gamma^{*i}_{kl} v^k du^l.$$

d^* deutet an, daß das Differential bezüglich des Zuwachses $(u+du, \dot{u}+d^*\dot{u})$ zu bilden ist. Um von dem Linienelement $(u+du, \dot{u}+d^*\dot{u})$ zu dem Linienelement von gleichem Zentrum $(u+du, \dot{u}+d\dot{u})$ zu gelangen, muß der Richtung $\dot{u}+d^*\dot{u}$ der Zuwachs

(2, 11) $$d\dot{u} - d^*\dot{u} = \overline{d}\dot{u}$$

erteilt werden. Dann kommt für $D_\varrho v^i$ wegen (1, 27)

(2, 12) $$D_\varrho v^i = \overline{d} v^i + C^i_{kl} v.$$

Die Summe der Differentiale $\overline{d} + d^*$ wollen wir mit d bezeichnen, dann erhalten wir aus (2, 9) zufolge (2, 10), (2, 11) und (2, 12) für Dv^i schließlich

(2, 13) $$Dv^i = dv^i + C^i_{kl} v^k d\dot{u}^l + \Gamma^{*i}_{kl} v^k du^l - C^i_{kl} v^k d^*\dot{u}^l.$$

In (2, 13) sind die $d^*\dot{u}^l$ diejenigen Differentiale, die das zu $\dot{u}^l$ parallele Linienelement im Punkte $(u+du)$ bestimmen. Für dieselben gilt aber (2, 6). Führt man dies in (2, 13) ein, so erhält man

(2, 14) $$Dv^i = dv^i + C^i_{kl} v^k d\dot{u}^l + \Gamma^i_{kl} v^k du^l,$$

wobei

(2, 15) $$\Gamma^i_{kl} = \Gamma^{*i}_{kl} + C^i_{ks} \Gamma^{*s}_{rl} u^r.$$

Da die Γ^{*i}_{kl} die zu dem Maßtensor γ_{ik} gehörigen Christoffelschen Symbole sind, erhält man bei Beachtung von (2, 5), (2, 2) und (2, 6)

(2, 16) $$\Gamma^*_{ilk} = \frac{1}{2}\left(\frac{\partial g_{il}}{\partial u^k} + \frac{\partial g_{kl}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^l}\right) - \\ - C_{ils} \Gamma^{*s}_{kj} \dot{u}^j - C_{kls} \Gamma^{*s}_{ij} \dot{u}^j + C_{iks} \Gamma^{*s}_{lj} u^j.$$

Überschiebt man (2, 16) mit $\dot{u}^i$ und $\dot{u}^k$, so ergibt sich wegen (1, 28), (2, 2) und der Homogenitätseigenschaft von $\mathsf{L}(u, \dot{u})$

(2, 17) $$\Gamma^*_{ilk} \dot{u}^i u^k = 2G_l = 2G^s g_{sl}.$$

Dabei haben wir

(2, 18) $$G_l = \frac{\partial^2(\mathsf{L}^2)}{\partial \dot{u}^l \partial u^k} \dot{u}^k - \frac{\partial(\mathsf{L}^2)}{\partial u^l}$$

gesetzt. Benützen wir die Gleichungen (2, 16) bis (2, 18), so erhält man

$$(2,19)\qquad \Gamma^{*s}_{kj}\,\dot u^k = \frac{\partial G^s}{\partial \dot u^j}.$$

Substituiert man dies in die linke Seite von (2, 16), so kommt schließlich

$$(2,20)\qquad \begin{aligned}\Gamma^*_{ilk} = \frac{1}{2}\left(\frac{\partial g_{il}}{\partial u^k} + \frac{\partial g_{lk}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^l}\right) - \\ - C_{ils}\frac{\partial G^s}{\partial \dot u^k} - C_{kls}\frac{\partial G^s}{\partial \dot u^i} + C_{iks}\frac{\partial G^s}{\partial \dot u^l}.\end{aligned}$$

Die Γ^*_{ilk} bestimmen zufolge (2, 15) auch die Übertragungsparameter Γ_{ilk} in der Form[8])

$$(2,21)\quad \Gamma_{ilk} = \frac{1}{2}\left(\frac{\partial g_{il}}{\partial u^k} + \frac{\partial g_{lk}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^l}\right) - C_{kls}\frac{\partial G^s}{\partial \dot u^i} + C_{iks}\frac{\partial G^s}{\partial \dot u^l}.$$

In der Finslerschen Geometrie[9]) wird das Bogenelement ebenfalls durch einen Ausdruck der Form (1, 4) bestimmt. Die „Grundfunktion“ $L(u, \dot u)$ genügt dabei denselben Voraussetzungen, wie im Falle einer Minkowskischen Geometrie, nur gibt es hier kein ausgezeichnetes Koordinatensystem, d. h. es ist unmöglich, durch eine Koordinatentransformation die Grundfunktion auf die Form (1, 1) zu bringen. Das invariante Differential für Finslersche Räume stimmt formal ebenfalls mit demjenigen in Minkowskischen Räumen überein, vorausgesetzt, daß derselbe auf krummlinige Koordinaten bezogen wird. Die Übertragungsparameter C_{ikl}, Γ^*_{ikl} und Γ_{ikl} drücken sich in beiden Fällen (immer unter der Voraussetzungen, daß der Minkowskische Raum auf krummlinige Koordinaten bezogen wird) genau auf dieselbe Weise durch den Fundamentaltensor und seine Ableitungen aus[10]). Wegen (1, 2) kann man auch sagen, sie bestimmen sich auf dieselbe Weise aus der Grundfunktion heraus. Da durch Bogenelement und invariantes Differential die Finslersche Geometrie im wesentlichen bestimmt ist, ist zu erwarten, daß diese Analogie sicher über das Formale hinausgeht. In der Tat läßt sich — wie schon einleitend bemerkt — zeigen, daß man den Finslerschen Raum mit Hilfe Minkowskischer Räume aufbauen kann.

---

8) Die Berechnung von (2, 20) und (2, 21) findet sich in etwas anderem Zusammenhange bei E. Cartan (2), S. 16.

9) Vgl. P. Finsler (1), E. Cartan (2) und O. Varga (1).

10) Vgl. O. Varga (1), S. 174.

§. 3. Die Krümmungstheorie.

Die Methoden zur Darstellung der Torsions- und Krümmungstheorie von Minkowskischen Räumen, die auf krummlinige Koordinaten bezogen sind, stimmen ganz mit denjenigen für Finslersche Räume überein. Dies folgt unmittelbar daraus, daß man in beiden Fällen wohl am einfachsten mit Hilfe der Cartanschen Methode der äußeren Produkte und äußeren Ableitung Pfaffscher Formen zu dieser Darstellung gelangt[11]). Dabei benützt man bei dieser Darstellung als Ausgangselemente die Tensoren g_{ik}, C_{ikl}, und ferner diejenigen, die sich aus ihnen durch Anwendung des invarianten Differentials herleiten lassen.

Die Torsionstensoren erhält man in dem Koeffizientensystem der äußeren quadratischen Form

(3, 1) $$\Omega^i = [du^s\, \omega(d)_s^i]$$

und die Krümmungstensoren aus dem Koeffizientensystem der Form

(3, 2) $$\Omega_k^i = [\omega_k^s(d)\, \omega_s^i(d)] - (\omega(d)_k^i)'.$$

Dabei wurde gesetzt

(3, 3) $$\omega(d)_s^i = \mathsf{L} \,.\, C_{sk}^{\cdot i}\, \omega^k(d) + \Gamma_{sk}^{*i} du^k$$

und $\omega^k(d)$ bedeutet das invariante Differential des Einheitsvektors l^i, also:

(3, 4) $$\omega^i(d) = Dl^i = D\left(\frac{u^i}{\mathsf{L}(u, u)}\right).$$

Da also die verwendeten Größen in beiden Geometrien formal übereinstimmen, erhält man vom formalen Standpunkte aus auch dieselben Torsions- und Krümmungstensoren[12]).

Wie in Finslerschen Räumen folgt aus (3, 1), daß auch der Minkowskische Raum nur einen Torsionstensor besitzt und zwar:

(3, 5) $$A_{ikl} = \mathsf{L}\, C_{ikl}.$$

Wegen der Koordinateninvarianz der ausgeführten Operationen können wir statt krummliniger Koordinaten auch die kartesischen Koordinaten des §. 1 verwenden und auf diese die Cartansche Symbolik anwenden. Da in den kartesischen Koordinaten die Γ_{ikl}^* verschwinden, zeigt sich, daß von den drei Krümmungstensoren,

---

[11]) E. Cartan (1), Kap. VII.

[12]) Wir verweisen hier so wie für den ganzen §. 3. auf E. Cartan (2), Kap. XIII. Die dort verwendete Bezeichnung der Krümmungstensoren wurde von uns übernommen.

die durch (3, 2) bestimmt werden, nur ein einziger von Null verschieden ist. Derselbe hat die Gestalt

$$(3,6) \qquad S_{ijkh} = L^2(u, \dot{u})\,(C_{imh}C^m_{jk} - C_{imk}C^m_{jh}).$$

P_{ijkh} und R_{ijkh} sind Nulltensoren[11]). Bei Benützung von kartesischen Koordinaten kann man weiter feststellen, daß auch der Tensor $C_{ikl|m}$, mit dessen Hilfe sich P_{ijkh} gewinnen läßt, verschwindet[13]).

Umgekehrt läßt sich nun zeigen, daß das Verschwinden von R_{ijkl} und $C_{ikl|m}$ die Minkowskischen Räume unter den Finslerschen auszeichnet. Diese Charakterisierung findet sich ohne Beweisangabe bei E. CARTAN[11]). Der Beweis kann etwa so geführt werden: Wir halten in $g_{ik}(u, \dot{u})$ $\dot{u}^i$ fest — betrachten diese Größen im wesentlichen als Parameter — und leiten dann aus diesen nunmehr nur von u^i abhängenden Tensor den vierfachen kovarianten Riemann—Christoffelschen Krümmungstensor her. Aus dem Verschwinden von R_{ijkl} und $C_{ikl|m}$ kann man auf Grund einer Rechnung schließen, daß auch der Riemann—Christoffelsche Krümmungstensor verschwindet. Dies bedeutet, wie aus der Theorie der Riemannschen Räume bekannt ist[14]), daß man die quadratische Form

$$(3,7) \qquad ds^2 = g_{ik}(u, \dot{u})\,du^i du^k$$

durch eine Koordinatentransformation auf eine Gestalt bringen kann, in der die Koeffizienten Konstante sind, die in unserem Falle aber natürlich von den Parametern $\dot{u}^i$ abhängen. Ist also

$$(3,8) \qquad u^i = u^i(x^1, x^2, \ldots, x^n) \qquad (i = 1, 2, \ldots, n)$$

diese Transformation, so erhält man für das Bogenelement

$$(3,9) \qquad ds^2 = \bar{g}_{ik}(\dot{x})\,dx^i dx^k.$$

Da die $\bar{g}_{ik}(\dot{x})$ nur von den $\dot{x}^i$ abhängen, folgt aus (2, 2), daß auch $L(x, \dot{x})$ nur von den $\dot{x}$ abhängt, also die Form (1, 1) hat, w. z. b. w. Die Koordinaten $x^1, x^2, \ldots, x^n$ sind dann die in §. 1 verwendeten kartesischen Koordinaten.

---

13) Vgl. E. CARTAN (2), S. 12. Die dort angegebene Formel XIV angewendet auf C_{ijk} gibt

$$(\mathrm{a}) \qquad C_{ikl|m} = \frac{\partial C_{ikl}}{\partial x^m} - \frac{\partial C_{ikl}}{\partial \dot{x}^s}\,\frac{\partial G^s}{\partial \dot{x}^m} - C_{skl}\Gamma^{*s}_{im} - C_{isl}\Gamma^{*s}_{km} - C_{iks}\Gamma^{*s}_{lm}.$$

Die linke Seite von (a) verschwindet, da C_{ikl} x^i nicht enthält und Γ^{*s}_{im} und daher auch $\dfrac{\partial G^s}{\partial \dot{x}^m}$ Null sind.

14) F. CARTAN (1), Kap. 2, insbes. S. 49—59.

Schriftenverzeichnis.

E. Cartan, (1) *Leçons sur la géométrie des espaces de Riemann* (Paris, 1928).

(2) Les espaces de Finsler, *Actualités scientifiques et industrielles*, 79 (Paris, 1934).

P. Finsler, (1) *Über Kurven und Flächen in allgemeinen Räumen* (Dissertation, Göttingen, 1918).

O. Varga, (1) Zur Herleitung des invarianten Differentials in Finslerschen Räumen, *Monatshefte für Math. und Phys.*, 50 (1941), S. 165–175.

(2) A Finsler-féle geometria felépítése Minkowski-féle simuló mértékmeghatározással (Aufbau der Finslerschen Geometrie mit Hilfe einer oskulierenden Minkowskischen Maßbestimmung), *Matematikai és Természettudományi Értesítő (Sitzungsberichte der ungarischen Akademie der Wissenschaften)*, 61 (1942), S. 14–22.

Einige Extremaleigenschaften des Kreisbogens bezüglich der Annäherung durch Polygone.

Von LÁSZLÓ FEJES in Kolozsvár.

Um das Maß der Approximation einer konvexen Kurve durch Polygonfolgen deuten zu können, muß man irgendeine Definition der „Abweichung" von zwei konvexen Kurven $\mathfrak{K}$ und $\mathfrak{L}$ zu Grunde legen. Eine in der Literatur gebrauchliche Definition ist die Folgende[1]: $d=\eta(\mathfrak{K}, \mathfrak{L})$ heiße *Streckenabweichung* zweier konvexen Kurven $\mathfrak{K}$ und $\mathfrak{L}$, wenn es die kleinste Zahl ist mit der Eigenschaft, daß der kürzeste Abstand jedes Punktes der Kurve $\mathfrak{K}$ von $\mathfrak{L}$ $\leqq d$ ist und umgekehrt der kürzeste Abstand jedes auf $\mathfrak{L}$ liegenden Punktes von $\mathfrak{K}$ ebenso $\leqq d$ ist.

Wir können ferner erklären: Betrachten wir den Flächeninhalt $\tau(\mathfrak{K}, \mathfrak{L})$ der Menge derjenigen Punkte, die von den von $\mathfrak{K}$ und $\mathfrak{L}$ begrenzten Gebieten genau in einem liegen. $\tau(\mathfrak{K}, \mathfrak{L})$ heiße *Inhaltsabweichung* zwischen $\mathfrak{K}$ und $\mathfrak{L}$.

Es sei bemerkt, daß wir $\tau(\mathfrak{K}, \mathfrak{L})$ auch als Differenz der Inhalte der Vereinigungsmenge und der Durchschnittsmenge der von $\mathfrak{K}$ und $\mathfrak{L}$ begrenzten Gebiete deuten können. Ganz analog nennen wir *Umfangsabweichung* die Differenz zwischen den Bogenlängen der Ränder der Vereinigungsmenge und der Durchschnittsmenge der von $\mathfrak{K}$ und $\mathfrak{L}$ begrenzten Gebiete und bezeichnen sie mit $\lambda(\mathfrak{K}, \mathfrak{L})$.

Die soeben definierten Abweichungen erfüllen die üblichen Forderungen: 1. Sie sind nichtnegativ und nur dann 0, wenn $\mathfrak{K}$

[1]) Vgl. BONNESEN—FENCHEL, *Theorie der konvexen Körper* (Berlin, 1934). Die obige Bezeichnung „Streckenabweichung" führe ich zwecks Unterscheidung von den im Folgenden einzuführenden Abweichungsdefinitionen ein.

und $\mathfrak{L}$ identisch sind. 2. Sie sind symmetrisch in beiden Kurven. 3. Sie genügen der „Dreiecksungleichung".

Betrachten wir nun eine vorgegebene konvexe Kurve $\mathfrak{K}$ und fassen das n-Eck $\mathfrak{P}_n$ mit der kleinstmöglichen Streckenabweichung $\eta(\mathfrak{K}, \mathfrak{P}_n)$ ins Auge. Wir erklären als *Maß der Approximierbarkeit* der Kurve $\mathfrak{K}$ durch Polygonfolgen den Grenzwert $\overline{\lim}\, n^2 \eta(\mathfrak{K}, \mathfrak{P}_n)$. Ganz analog können wir das Approximierbarkeitsmaß mit Hilfe der Inhalts-, oder Umfangsabweichung deuten. Wir sagen ferner, daß sich $\mathfrak{K}$ bezüglich einer der drei Abweichungen schlechter, gleich gut, oder besser approximieren läßt als $\mathfrak{L}$, je nachdem das Verhältnis ihrer Approximierbarkeitsmaße > 1, $= 1$, oder < 1 ausfällt.

Es erhebt sich nun eine Reihe interessanter Probleme, wenn wir unter Kurven, die gewissen Bedingungen unterworfen sind, diejenige bestimmen wollen, die sich in obigem Sinne am schlechtesten durch Polygonfolgen approximieren läßt.

Die Lösung eines derartigen Problems gibt z. B. folgender Satz[2]: Unter sämtlichen geschlossenen konvexen Kurven mit vorgegebenem Flächeninhalt läßt sich bezüglich der Inhaltsabweichung die Ellipse am schlechtesten durch Polygonfolgen approximieren.

Dasselbe gilt auch für Approximation durch einbeschriebene[3], sowie durch umbeschriebene[4] Polygonfolgen.

Wegen der isoperimetrischen Ungleichung folgt hieraus, daß für Kurven mit vorgegebener Bogenlänge der Kreis die extremale Kurve ist. Im Folgenden nehmen wir an, daß die zur Konkurrenz zugelassenen Kurven bis auf eine endliche Anzahl von Punkten stetig veränderliche Krümmung besitzen[5] und zeigen, daß unter dieser Voraussetzung für Kurven mit vorgegebener Bogenlänge stets der Kreis die extremale Kurve ist, gleichgültig welche der obigen Abweichungsdefinitionen wir zu Grunde legen.

---

[2]) L. Fejes, Sur un théorème concernant l'approximation des courbes par des suites de polygones, *Annali di Pisa*, (2) **9** (1940), S. 143—145.

[3]) E. Sas, Über eine Extremumeigenschaft der Ellipsen, *Compositio Math.*, **6** (1939), S. 468-470.

[4]) L. Fejes, Eine Bemerkung zur Approximation durch n-Eckringe, *Compositio Math.*, **7** (1940), S. 474—476.

[5]) Unsere Ergebnisse gelten vermutlich auch ohne diese Einschränkung.

Es gilt noch etwas allgemeiner der Satz: *Unter sämtlichen konvexen stetig gekrümmten Kurvenbogen mit vorgegebener Bogenlänge und totaler Biegung*[6]) *läßt sich bezüglich der Strecken-, der Inhalts- und auch der Umfangsabweichung der Kreisbogen am schlechtesten durch Streckenzugfolgen approximieren. Dasselbe gilt auch für eine Folge einbeschriebener, sowie umbeschriebener Streckenzüge*[7]).

In diesem Satz sind neun verschiedene Ungleichungen enthalten. Zum Beweis stellen wir zunächst die neun Approximierbarkeitsmaße für einen Kreis zusammen.

Wir beginnen mit dem einbeschriebenen n-Eck. Unter sämtlichen einem Kreise $\mathfrak{K}$ mit dem Halbmesser r einbeschriebenen n-Ecken hat offenbar das reguläre n-Eck $\mathfrak{P}_n^e$ die kleinstmöglichen Abweichungen von $\mathfrak{K}$. Es gilt:

$$\eta(\mathfrak{K}, \mathfrak{P}_n^e) = r\left(1 - \cos\frac{\pi}{n}\right), \quad \tau(\mathfrak{K}, \mathfrak{P}_n^e) = \frac{nr^2}{2}\left(\frac{2\pi}{n} - \sin\frac{2\pi}{n}\right),$$

$$\lambda(\mathfrak{K}, \mathfrak{P}_n^e) = 2nr\left(\frac{\pi}{n} - \sin\frac{\pi}{n}\right),$$

woraus sich

$$\lim n^2\eta(\mathfrak{K}, \mathfrak{P}_n^e) = \frac{r\pi}{2}, \quad \lim n^2\tau(\mathfrak{K}, \mathfrak{P}_n^e) = \frac{2r^2\pi^3}{3},$$

$$\lim n^2\lambda(\mathfrak{K}, \mathfrak{P}_n^e) = \frac{r\pi^3}{3}$$

ergibt.

Für umbeschriebene n-Ecke hat ebenfalls das reguläre n-Eck $\mathfrak{P}_n^u$ die kleinstmöglichen Abweichungen vom Kreis:

$$\eta(\mathfrak{K}, \mathfrak{P}_n^u) = r\,\frac{1-\cos\frac{\pi}{n}}{\cos\frac{\pi}{n}}, \quad \tau(\mathfrak{K}, \mathfrak{P}_n^u) = nr^2\left(\operatorname{tg}\frac{\pi}{n} - \frac{\pi}{n}\right),$$

$$\lambda(\mathfrak{K}, \mathfrak{P}_n^u) = 2nr\left(\operatorname{tg}\frac{\pi}{n} - \frac{\pi}{n}\right).$$

---

[6]) Ist $\varrho = \varrho(s)$ die „natürliche Gleichung" eines konvexen Kurvenbogens $\mathfrak{b}$, so heißt $\omega = \int_0^L \varrho(s)\,ds$ totale Biegung von $\mathfrak{b}$.

[7]) Um die Abweichung zwischen einem konvexen Kurvenbogen $\mathfrak{b}$ und einem Streckenzug $\mathfrak{p}_n$ — und dadurch die Approximierbarkeit — nach der obigen Erklärung deuten zu können, denken wir uns $\mathfrak{b}$ und $\mathfrak{p}_n$ durch die zu den Endpunkten gehörigen Sehnen stets zu geschloßenen Kurven ergänzt.

Es gilt daher:

$$\lim n^2\eta(\mathfrak{K}, \mathfrak{P}_n^\eta) = \frac{r\pi^2}{2}, \quad \lim n^2\tau(\mathfrak{K}, \mathfrak{P}_n^\tau) = \frac{r^2\pi^3}{3},$$
$$\lim n^2\lambda(\mathfrak{K}, \mathfrak{P}_n^\lambda) = \frac{2r\pi^3}{3}.$$

Betrachten wir nun den Fall der beliebigen n-Ecke! Hier fallen die drei n-Ecke mit den kleinsten Abweichungen nicht zusammen. Bei der Streckenabweichung ist es ein mit $\mathfrak{K}$ konzentrisches reguläres n-Eck, das dadurch ausgezeichnet ist, daß das aritmetische Mittel der Halbmesser seiner In- und Umkreise gleich r ist. Für dieses n-Eck $\mathfrak{P}_n$ gilt:

$$\eta(\mathfrak{K}, \mathfrak{P}_n) = r\,\frac{1-\cos\frac{\pi}{n}}{1+\cos\frac{\pi}{n}}, \quad \lim n^2\eta(\mathfrak{K}, \mathfrak{P}_n) = \frac{r\pi^2}{4}.$$

Das n-Eck mit der kleinstmöglichen Inhaltsabweichung ist wiederum ein mit $\mathfrak{K}$ konzentrisches reguläres n-Eck mit der Eigenschaft, daß die außerhalb und innerhalb von $\mathfrak{K}$ liegende Teile seines Umfanges gleich sind. Für die Inhaltsabweichung selbst läßt sich hierbei keine so durchsichtige Formel angeben wie die obigen. Man rechnet jedoch leicht nach, daß

$$\lim n^2\tau(\mathfrak{K}, \mathfrak{P}_n) = \frac{r^2\pi^3}{4}$$

ausfällt.

Es fehlt nun noch die Berechnung des Approximierbarkeitsmaßes bezüglich der Umfangsabweichung. Hier sei die Frage der Bestimmung des extremalen n-Eckes $\mathfrak{P}_n$ beiseite gelassen. Wir machen nur von der trivialen Tatsache Gebrauch, daß jede Seite von $\mathfrak{P}_n$ zwei gemeinsame Punkte mit $\mathfrak{K}$ besitzt.

Greifen wir eine Seite von $\mathfrak{P}_n$, etwa die i-te $\overline{A_iA_{i+1}}$, heraus. Ihre Schnittpunkte mit $\mathfrak{K}$ seien B_i und C_i. Es gilt

$$\lambda(\mathfrak{K}, \mathfrak{P}_n) = \sum_{i=1}^{n} [\overline{A_iA_{i+1}} + 2(\widehat{B_iC_i} - \overline{B_iC_i})] - 2\pi r.$$

Setzen wir $\sphericalangle A_iOA_{i+1} = \alpha_i$, $\sphericalangle B_iOC_i = \beta_i$, wobei O der Kreismittelpunkt ist, so gilt

$$\overline{A_iA_{i+1}} \geqq 2r\cos\frac{\beta_i}{2}\,\mathrm{tg}\,\frac{\alpha_i}{2}.$$

Betrachten wir nähmlich die Strecke $\overline{A_i'A_{i+1}'}$, welche die Schenkel des Winkels $\sphericalangle A_i'OA_{i+1}' = \alpha_i$ aus der fest gedachten Geraden A_iA_{i+1} ausschneiden. Drehen wir das Schenkelpaar bei festem α_i, so wird das Minimum von $\overline{A_i'A_{i+1}'}$ im Falle $\overline{OA_i'} = \overline{OA_{i+1}'}$ erreicht. Es gilt ferner

$$\widehat{B_iC_i} - \overline{B_iC_i} = r\left(\beta_i - 2\sin\frac{\beta_i}{2}\right)$$

und daher

$$\lambda(\mathfrak{K}, \mathfrak{P}_n) \geqq 2r\sum_{i=1}^{n}\left(\cos\frac{\beta_i}{2}\operatorname{tg}\frac{\alpha_i}{2} + \beta_i - 2\sin\frac{\beta_i}{2}\right) - 2\pi r.$$

Beachten wir nun, daß $\beta_i \leqq \alpha_i$ und $(\alpha_i)_{\max} \to 0$, so erhalten wir

$$\cos\frac{\beta_i}{2}\operatorname{tg}\frac{\alpha_i}{2} + \beta_i - 2\sin\frac{\beta_i}{2} = \left(1 - \frac{\beta_i^2}{8}\right)\left(\frac{\alpha_i}{2} + \frac{\alpha_i^3}{24}\right) + \frac{\beta_i^3}{24} + O(\alpha_i^5) =$$

$$= \frac{\alpha_i}{2} + \frac{1}{8}\left(\frac{\alpha_i^3}{3} + \frac{\beta_i^3}{3} - \frac{\alpha_i\beta_i^2}{2}\right) + O(\alpha_i^5).$$

Die Differenz $\frac{\beta_i^3}{3} - \frac{\alpha_i\beta_i^2}{2}$ erreicht aber für $0 \leqq \beta_i \leqq \alpha_i$ ihr Minimum im Falle $\beta_i = \alpha_i$, wodurch sich

$$\cos\frac{\beta_i}{2}\operatorname{tg}\frac{\alpha_i}{2} + \beta_i - 2\sin\frac{\beta_i}{2} \geqq \frac{\alpha_i}{2} + \frac{\alpha_i^3}{48} + O(\alpha_i^5)$$

ergibt. Man beachte nun, daß $\sum_{i=1}^{n} \alpha_i = 2\pi$ ist. Unter dieser Bedingung nimmt aber $\sum_{i=1}^{n} \alpha_i^3$ — laut der Jensenschen Ungleichheit — ihr Minimum für $\alpha_i = \frac{2\pi}{n}$ $(i = 1, 2, \ldots, n)$ an: $\sum_{i=1}^{n} \alpha_i^3 \geqq \frac{8\pi^3}{n^2}$. Es gilt daher für das Approximierbarkeitsmaß

$$\lim n^2 \lambda(\mathfrak{K}, \mathfrak{P}_n) = \frac{r\pi^3}{3}.$$

Folglich erhalten wir die überraschende Tatsache, daß der Kreis bezüglich der Umfangsabweichung nicht besser durch eine Folge beliebiger Polygone approximiert werden kann als durch eine einbeschriebene Polygonfolge.

Aus den Obigen erhalten wir die Approximierbarkeitsmaße für einen Kreisbogen $\mathfrak{b}$ mit der Bogenlänge L und totaler Biegung ω durch einbeschriebene, umbeschriebene oder beliebige Streckenzugfolgen $\{\mathfrak{p}_n^e\}$, $\{\mathfrak{p}_n^u\}$ bzw. $\{\mathfrak{p}_n\}$. Wir haben nur in unseren Formeln

2π durch ω und r durch $\frac{L}{\omega}$ zu ersetzen:

$$\lim n^2 \eta(\mathfrak{b}, \mathfrak{p}_n^e) = \frac{1}{8} L\omega, \quad \lim n^2 \eta(\mathfrak{b}, \mathfrak{p}_n^u) = \frac{1}{8} L\omega,$$

$$\lim n^2 \eta(\mathfrak{b}, \mathfrak{p}_n) = \frac{1}{16} L\omega,$$

$$\lim n^2 \tau(\mathfrak{b}, \mathfrak{p}_n^e) = \frac{1}{12} L^2\omega, \quad \lim n^2 \tau(\mathfrak{b}, \mathfrak{p}_n^u) = \frac{1}{24} L^2\omega,$$

$$\lim n^2 \tau(\mathfrak{b}, \mathfrak{p}_n) = \frac{1}{32} L^2\omega,$$

$$\lim n^2 \lambda(\mathfrak{b}, \mathfrak{p}_n^e) = \frac{1}{24} L\omega^2, \quad \lim n^2 \lambda(\mathfrak{b}, \mathfrak{p}_n^u) = \frac{1}{12} L\omega^2,$$

$$\lim n^2 \lambda(\mathfrak{b}, \mathfrak{p}_n) = \frac{1}{24} L\omega^2.$$

Nun zum Beweis unseres Satzes! Wir beweisen ihn zunächst für einen „Kreisbogenzug", d. h. für einen Kurvenbogen, der aus einer endlichen Anzahl von Kreisbogen besteht. Es sei zunächst bemerkt, daß es genügt den Satz für Kreisbogenzüge aus zwei Kreisbogen zu beweisen. Ersetzt man nämlich zwei anstoßende Kreisbogen stets durch einen einzigen derart, daß Bogenlänge und totale Biegung unverändert bleiben, so erhält man schließlich einen einzigen Kreisbogen, der sich im Verhältnis zum Kreisbogenzug schlechter approximieren läßt.

Betrachten wir demnach einen Kreisbogenzug $\mathfrak{b}$ aus zwei Kreisbogen $\mathfrak{b}_1$ und $\mathfrak{b}_2$. Es gilt $L = L_1 + L_2$, $\omega = \omega_1 + \omega_2$, wobei L, L_1, L_2 die entsprechenden Bogenlängen, ω, ω_1, ω_2 die totalen Biegungen bedeuten.

Betrachten wir zunächst die Approximation bezüglich der Streckenabweichung. Approximieren wir $\mathfrak{b}$ durch einen Streckenzug $\mathfrak{p}_n$ mit n Seiten, wobei die zu $\mathfrak{b}_1$ und $\mathfrak{b}_2$ gehörigen Seitenzahlen n_1 und n_2 $(n_1 + n_2 = n)$ derart gewählt werden sollen, daß für $n \to \infty$ $\frac{n_1^2}{n_2^2} \to \frac{L_1\omega_1}{L_2\omega_2}$ sei. Das Approximierbarkeitsmaß von $\mathfrak{b}$ ist — abgesehen vom Faktor $\frac{1}{8}$ bzw. $\frac{1}{16}$, der hinzuzufügen ist, je nachdem die Annäherung durch ein- oder umbeschriebene, bzw. beliebige Streckenzüge erfolgt —

$$\lim n^2 \frac{L_1\omega_1}{n_1^2} = \lim n^2 \frac{L_2\omega_2}{n_2^2}.$$

Wir zeigen, daß dasselbe $\leqq L\omega$ ausfällt. Es gilt nämlich

$$\lim n^2 \frac{L_1\omega_1}{n_1^2} = \lim\left(\frac{n_1+n_2}{n_1}\right)^2 L_1\omega_1 =$$

$$= \left(1+\sqrt{\frac{L_2\omega_2}{L_1\omega_1}}\right)^2 L_1\omega_1 = (\sqrt{L_1\omega_1}+\sqrt{L_2\omega_2})^2 \leqq$$

$$\leqq (\sqrt{L_1\omega_1}+\sqrt{L_2\omega_2})^2 + (\sqrt{L_1\omega_2}-\sqrt{L_2\omega_1})^2 = (L_1+L_2)(\omega_1+\omega_2) = L\omega.$$

Das Gleichheitszeichen gilt nur im Falle $(\sqrt{L_1\omega_2}-\sqrt{L_2\omega_1})^2 = 0$, also $\frac{L_1}{\omega_1} = \frac{L_2}{\omega_2}$, d. h. falls der Kreisbogenzug in einem einzigen Kreisbogen entartet.

Wenden wir uns nun der Approximation bezüglich der Inhaltsabweichung zu. Hier wählen wir die Seitenzahlfolgen $\{n_1\}$ und $\{n_2\}$ so, daß $\frac{n_1^3}{n_2^3} \to \frac{L_1^2\omega_1}{L_2^2\omega_2}$ sei. Wir zeigen, daß dann die Ungleichung

$$\lim n^2\left(\frac{L_1^2\omega_1}{n_1^2}+\frac{L_2^2\omega_2}{n_2^2}\right) \leqq L^2\omega$$

gilt, womit alles bewiesen sein wird. Um sie einzusehen, beachten wir, daß

$$\lim n^2\left(\frac{L_1^2\omega_1}{n_1^2}+\frac{L_2^2\omega_2}{n_2^2}\right) = (\sqrt[3]{L_1^2\omega_1}+\sqrt[3]{L_2^2\omega_2})^3$$

ist. Setzen wir

$$F(\omega_1, \omega_2) = \sqrt[3]{L_1^2\omega_1}+\sqrt[3]{L_2^2\omega_2}$$

und stellen die Aufgabe: das Maximum von $F(\omega_1, \omega_2)$ unter der Bedingung $\omega_1+\omega_2=\omega$ und festem L_1 und L_2 zu bestimmen. Dazu dient die Gleichung

$$\frac{dF}{d\omega_1} = \frac{\partial F}{\partial\omega_1} + \frac{\partial F}{\partial\omega_2}\frac{d\omega_2}{d\omega_2} = \frac{1}{3}\left(\sqrt[3]{\frac{L_1^2}{\omega_1^2}} - \sqrt[3]{\frac{L_2^2}{\omega_2^2}}\right) = 0,$$

welche nur für $\frac{L_1}{\omega_1} = \frac{L_2}{\omega_2}$ bestehen kann, d. h. wiederum im Falle eines einzigen Kreisbogens. Es gilt ferner

$$\frac{d^2F}{d\omega_1^2} = -\frac{2}{9}\left(\sqrt[3]{\frac{L_1^2}{\omega_1^5}} + \sqrt[3]{\frac{L_2^2}{\omega_2^5}}\right).$$

Dies ist aber für positive Werte von ω_1 und ω_2 stets negativ, was in Evidenz setzt, daß von einem absoluten Maximum die Rede ist.

Aus $\omega_2 L_1 = \omega_1 L_2$, $L_1 + L_2 = L$ erhält man nun $L_1 = \frac{\omega_1}{\omega} L$, $L_2 = \frac{\omega_2}{\omega} L$, und mithin $\left(\sqrt[3]{L_1^2 \omega_2} + \sqrt[3]{L_2^2 \omega_2}\right)^3 = L^2 \omega$, womit unsere Ungleichung völlig bewiesen ist.

Im Falle der Approximation bezüglich der Umfangsabweichung verfährt man ganz analog. Man wählt die Folgen $\{n_1\}$, $\{n_2\}$ so, daß $\frac{n_1^3}{n_2^3} \to \frac{L_1 \omega_1^2}{L_2 \omega_2^2}$ sei und prüft die Ungleichung

$$\lim n^2 \left(\frac{L_1 \omega_1^2}{n_1^2} + \frac{L_2 \omega_2^2}{n_2^2} \right) = \left(\sqrt[3]{L_1 \omega_1^2} + \sqrt[3]{L_2 \omega_2^2} \right)^3 \leqq L \omega^2.$$

Bemerken wir jedoch, daß diese schon aus den Obigen folgt, wenn wir in der soeben bewiesenen Ungleichung $\left(\sqrt[3]{L_1^2 \omega_1} + \sqrt[3]{L_2^2 \omega_2}\right)^3 \leqq L^2 \omega$ die Größen L_1, L_2 mit ω_1, ω_2 wertauschen.

Damit ist unser Satz für Kreisbogenzüge erledigt. Der Beweis für stetig gekrümmte Kurven wird nunmehr keine Schwierigkeiten bieten. Er beruht auf der einleuchtenden Tatsache, daß ein stetig gekrümmter konvexer Kurvenbogen besser durch eine Folge von Kreisbogenzügen angenähert werden kann, als durch gewöhnliche Streckenzugfolgen. Genauer: es läßt sich zu jedem stetig gekrümmten konvexen Kurvenbogen $\mathfrak{b}$ eine Folge von einbeschriebenen konvexen Kreisbogenzügen $\{\mathfrak{b}_n\}$ mit den Seitenzahlen n derart angeben, daß

$$\overline{\lim}\, n^2 \eta(\mathfrak{b}, \mathfrak{b}_n) = \overline{\lim}\, n^2 \tau(\mathfrak{b}, \mathfrak{b}_n) = \overline{\lim}\, n^2 \lambda(\mathfrak{b}, \mathfrak{b}_n) = 0$$

ausfällt. Dasselbe gilt auch für umbeschriebene Kreisbogenzüge[8]).

Approximieren wir demnach $\mathfrak{b}$ durch eine Folge von kon-

---

[8]) Auf den Beweis dieser Tatsache — da sie im wesentlichen nur eine andere Formulierung elementarer differentialgeometrischen Überlegungen ist — wollen wir nicht näher eingehen. Es sei jedoch bemerkt, daß es genügt $\overline{\lim}\, n^2 \eta(\mathfrak{b}, \mathfrak{b}_n) = 0$ nachzuweisen. Betrachten wir statt des Kurvenbogens — was auf dasselbe hinauskommt — zwei geschlossene konvexe Kurven $\mathfrak{K}$ und $\mathfrak{L}$, von denen $\mathfrak{L}$ in $\mathfrak{K}$ liege. Es gilt $\tau(\mathfrak{K}, \mathfrak{L}) \leqq L \eta(\mathfrak{K}, \mathfrak{L}) + \pi \eta(\mathfrak{K}, \mathfrak{L})^2$, wobei L die Bogenlänge von $\mathfrak{L}$ bezeichnet. Dies leuchtet ein, wenn wir bedenken, daß auf der rechten Seite das Inhaltsmaß von $\mathfrak{L}$ und ihrer im Abstand $\eta(\mathfrak{K}, \mathfrak{L})$ laufenden äußeren Parallelkurve steht. Ähnlich sieht man die Ungleichheit $\lambda(\mathfrak{K}, \mathfrak{L}) \leqq 2\pi \eta(\mathfrak{K}, \mathfrak{L})$ ein.

vexen einbeschriebenen Kreisbogenzügen $\{\mathfrak{b}_{\nu_n}\}$. Dabei sei die Seitenzahlfolge $\{\nu_n\}$ derart gewählt, daß außer $\lim n^2\eta(\mathfrak{b}, \mathfrak{b}_{\nu_n}) = 0$ auch noch die Bedingung $\frac{\nu_n}{n} \to 0$ bestehe. Unsere vorigen Überlegungen zeigen, daß es zu jedem $\mathfrak{b}_{\nu_n}$ einen einbeschriebenen Streckenzug $\mathfrak{p}_n$ mit der Seitenzahl n gibt so, daß

$$\overline{\lim}\, n^2\eta(\mathfrak{b}_{\nu_n}, \mathfrak{p}_n) \leqq \frac{1}{8} L\omega$$

gilt[9]). Die Streckenzüge $\{\mathfrak{p}_n\}$ sind aber offenbar auch $\mathfrak{b}$ einbeschrieben und es gilt wegen

$$\eta(\mathfrak{b}, \mathfrak{p}_n) \leqq \eta(\mathfrak{b}, \mathfrak{b}_{\nu_n}) + \eta(\mathfrak{b}_{\nu_n}, \mathfrak{p}_n)$$

die Ungleichung

$$\overline{\lim}\, n^2\eta(\mathfrak{b}, \mathfrak{p}_n) \leqq \frac{1}{8} L\omega.$$

Ganz analog prüft man die übrigen acht Ungleichungen.

Es bleibt noch übrig, die Unizität nachzuweisen. Man zerlege den als extremal vorausgesetzten Bogen $\mathfrak{b}$ in zwei beliebige Teilbogen $\mathfrak{b}_1$ und $\mathfrak{b}_2$. Das Approximierbarkeitsmaß bleibt unverändert, wenn wir $\mathfrak{b}_1$ und $\mathfrak{b}_2$ durch zwei Kreisbogen mit bzw. denselben Bogenlängen L_1, L_2 und totalen Biegungen ω_1, ω_2 ersetzen. Nun muß aber $\frac{L_1}{\omega_1} = \frac{L_2}{\omega_2}$ gelten, und zwar für eine beliebige Zerlegung. Daraus folgt für den Krümmungshalbmesser in einem beliebigen Punkt: $\frac{dL}{d\omega} = \text{Const.}$, d. h. $\mathfrak{b}$ erweist sich in der Tat als Kreisbogen.

Zum Schluß wollen wir noch auf einige zur Zeit ungelöste Probleme aufmerksam machen. Es läßt sich zeigen, daß es eine universelle Konstante, etwa 500, gibt, so, daß das Approximierbarkeitsmaß bezüglich der Umfangsabweichung einer konvexen Kurve mit dem Flächeninhalt T nicht die Größe $500T^{1/2}$ überschreiten kann[10]). Dies gilt auch für die Approximation durch einbeschriebene wie auch durch umbeschriebene Polygone und

[9]) Dies haben wir nur für einen einzigen festen Kreisbogenzug gezeigt. Unsere Überlegungen lassen sich aber mit Hilfe einer geeigneten Modifikation ohne Mühe auch auf den angedeuteten Fall ausdehnen.

[10]) L. Fejes, Über die Approximation konvexer Kurven durch Polygonfolgen, *Compositio Math.*, 6 (1939), S. 456—467.

zwar ohne Einschränkungen über die Krümmung. Dasselbe kann man auch — mit Hilfe einer analogen Schlußweise — von der Approximation bezüglich der Streckenabweichung behaupten. Es liegt nun die Vermutung nahe, daß im Falle der Strecken-, sowie der Umfangsabweichung auch dann der Kreis die extremale Kurve liefert, wenn statt der Bogenlänge der Flächeninhalt vorgegeben wird.

Wir kommen schließlich zu einer Fülle hübscher Probleme, wenn wir unter Kurven, die gewissen Bedingungen unterworfen sind, nach derjenigen fragen, die sich durch ein n-Eck mit festem n bezüglich irgendeiner der obigen Abweichungen am schlechtesten annähern läßt.

Ein derartiges Problem scheint im Verhältnis zu den Vorigen recht schwierig zu sein. In dieser Richtung ist mir nur die von mir angeregte Untersuchung vom Herrn E. SAS bekannt[11]). Sein Ergebnis lautet: Unter der Gesamtheit der konvexen geschlossenen Kurven mit vorgegebenem Inhaltsmaß erreicht der größtmögliche Flächeninhalt der einbeschriebenen n-Ecke sein Minimum für die Ellipsen.

(Eingegangen am 11. November 1941.)

---

[11]) S. Fußnote [3]).

Über konvexe Kurven und einschließende Kreisringe.

Von GYULA (JULIUS) v. Sz. NAGY in Kolozsvár.

Als Kreisring wird die abgeschlossene Punktmenge zwischen zwei konzentrischen Kreisen bezeichnet. Derjenige Kreisring mit minimaler Breite (mit minimaler Halbmesserdifferenz), der eine konvexe Kurve C enthält, heißt *Minimalkreisring*, wenn sein innerer Kreis innerhalb C liegt. Ein Kreisring, der C enthält und den kleinsten Flächeninhalt besitzt, soll als *kleinster Kreisring* von C bezeichnet werden, wenn der innere Kreis des Kreisringes innerhalb C liegt.

T. BONNESEN und N. KRITIKOS[1]) haben gezeigt, daß der Minimalkreisring einer konvexen Kurve C durch die Kurve eindeutig bestimmt wird.

Wir beweisen den folgenden Satz:

Eine konvexe Kurve hat im allgemeinen nur einen kleinsten Kreisring. Hat die konvexe Kurve C mehrere kleinste Kreisringe, so enthält sie eine Strecke, deren Endpunkte für C Ecken sind.

Eine konvexe Kurve mit diesen Eigenschaften muß aber nicht immer mehrere kleinste Kreisringe haben. Zum Rand des gleichseitigen Dreieckes gehört z. B. nur ein kleinster Kreisring.

1. Die Existenz mindestens eines kleinsten Kreisringes einer konvexen Kurve C erkennt man so:

---

[1]) Vgl. T. BONNESEN, Über das isoperimetrische Defizit ebener Figuren, *Math. Annalen*, **91** (1924), S. 252—268; N. KRITIKOS, Über konvexe Flächen und einschließende Kugeln, *Math. Annalen*, **96** (1927), S. 583—586; T. BONNESEN und W. FENCHEL, *Theorie der konvexen Körper* (Berlin, 1934), S. 54—55. Diese Arbeiten werden einfach unter BONNESEN, KRITIKOS bzw. BONNESEN—FENCHEL angeführt.

Es sei B der von C begrenzte abgeschlossene konvexe Bereich. Ist M ein Punkt von B, so gibt es einen kleinsten Kreis $K(M)$ mit dem Mittelpunkt M und einem Halbmesser $R(M)$, der B enthält, und einen größten Kreis $k(M)$ mit demselben Mittelpunkt und vom Halbmesser $r(M)$, der in B enthalten ist. Liegt M auf C, so reduziert sich $k(M)$ auf einen Punkt.

Der von $K(M)$ und $k(M)$ begrenzte Kreisring hat den Inhalt $\pi F(M)=\pi[R^2(M)-r^2(M)]$. Hier ist $F(M)$ in B eine nichtnegative, stetige Funktion von M. Nach einem Satz von WEIERSTRASZ erreicht die Funktion $F(M)$ in B ihr Minimum, das offenbar dann und nur dann gleich Null ist, wenn C ein Kreis ist.

Es ist klar, daß sowohl $K(M)$, wie $k(M)$ mit C mindestens einen Punkt gemeinsam haben.

2. Zum Beweis des ausgesprochenen Satzes benötigen wir den folgenden

Hilfssatz. *Wir bezeichnen mit k_i einen Kreis mit dem Mittelpunkt M_i und dem Halbmesser r_i ($i=1,2$ und $M_1 \neq M_2$), mit H die konvexe Hülle von k_1 und k_2, und mit Q einen außerhalb H liegenden Punkt der Potenzlinie von k_1 und k_2. In H gibt es dann und nur dann keinen solchen Kreis, dessen Mittelpunkt auf der Strecke M_1M_2 liegt und dessen Potenz bezüglich Q kleiner ist, als diejenige von k_1 und k_2, wenn einer der Kreise k_1 und k_2 den anderen von innen berührt.*

Wir nehmen an, daß der Halbmesser r_2 von k_2 nicht größer ist, als jener von k_1. Wir bezeichnen mit S den Schnittpunkt der Potenzlinie von k_1 und k_2 mit der Geraden M_1M_2.

Ist S ein innerer Punkt der Strecke $\overline{M_1M_2}$, so liegt der Kreis k mit dem Mittelpunkt S und vom Halbmesser r_2 offenbar in H und hat in bezug auf den Punkt Q eine kleinere Potenz, als k_2, weil $\overline{QS}^2-r_2^2<\overline{QM}_2^2-r_2^2$ ist.

Ist S ein innerer Punkt von H, aber kein innerer Punkt der Strecke $\overline{M_1M_2}$, so schneiden sich die Kreise k_1 und k_2 in zwei Punkten A und B. Derjenige Kreis k' durch die Punkte A und B, dessen Mittelpunkt M ein innerer Punkt der Strecke $\overline{M_1M_2}$ ist, hat keinen Punkt außerhalb von k_1 und k_2. Die konvexe Hülle von k_1 und k_2 enthält einen zu k' konzentrischen größeren Kreis k, weil A und B innere Punkte von H sind. Bezeichnet r' bzw. r den Halbmesser von k' bzw. k, so ist die Potenz von Q bezüglich

k, k' bzw. k_2

$$\overline{QM}^2 - r^2,\ \overline{QM}^2 - r'^2 \quad \text{bzw.} \quad \overline{QM}_2^2 - r_2^2.$$

Daraus folgt die Ungleichung

$$\overline{QM}^2 - r^2 < \overline{QM}^2 - r'^2 = \overline{QM}_2^2 - r_2^2,$$

weil $r > r'$ ist und weil Q auf der Potenzlinie AB der Kreise k' und k_2 liegt.

Liegt S außerhalb von H, so stimmt H mit k_1 überein. Bezeichnet S_1 den Schnittpunkt von k_1 mit der Strecke $\overline{SM_1}$ und k denjenigen Kreis vom Halbmesser r_2, der den Kreis k_1 in S_1 von innen berührt, dann ist die Potenz von Q bezüglich k kleiner als jene bezüglich k_2, weil $\overline{QM} < \overline{QM}_2$ ist, wo M den Mittelpunkt von k bedeutet.

Es bleibt nur der Fall übrig, in dem S auf den Rand von H fällt. Dann berührt k_2 den Kreis k_1 in S von innen. Der Punkt Q liegt dann auf der Tangente von k_1 in S. Die Potenz von Q in bezug auf k_1 und k_2 ist gleich $\overline{QS}^2$. Ist nun M ein beliebiger Punkt der Strecke $\overline{M_1M_2}$ und ist k ein in k_1 liegender Kreis mit dem Mittelpunkt M und vom Halbmesser r, so hat Q bezüglich k die Potenz

$$\overline{QM}^2 - r^2 \geqq \overline{QM}^2 - \overline{SM}^2 = \overline{QS}^2.$$

Damit ist der Hilfssatz vollständig bewiesen.

3. Wir nehmen nun an, daß die konvexe Kurve C mehrere kleinste Kreisringe hat und M_1 und M_2 die Mittelpunkte von zwei solchen Kreisringen sind.

Der konvexe Bereich B enthält die Kreise $k(M_1)$ und $k(M_2)$ und ihre konvexe Hülle H. Die Kreise $K(M_1)$ und $K(M_2)$ enthalten B und haben mindestens je einen Punkt mit C gemeinsam. Daraus folgt, daß $K(M_1)$ und $K(M_2)$ einander schneiden, oder der eine Kreis den anderen von innen berührt.

Schneiden $K(M_1)$ und $K(M_2)$ einander in Q und Q', so enthält ihr gemeinsames Kreiszweieck den Bereich B und damit auch H. Der Punkt Q liegt auf der Potenzlinie von $k(M_1)$ und $k(M_2)$, da

$$F(M_1) \equiv R^2(M_1) - r^2(M_1) = F(M_2) \equiv R^2(M_2) - r^2(M_2),$$
$$R(M_1) = \overline{QM}_1 \text{ und } R(M_2) = \overline{QM}_2.$$

Berührt nun keiner der Kreise $k(M_1)$ und $k(M_2)$ den anderen von innen, so gibt es nach dem Hilfssatz in H einen Kreis k, dessen Mittelpunkt M auf der Strecke M_1M_2 liegt und dessen Potenz bezüglich Q kleiner ist als die Potenz von $k(M_2)$ bezüglich Q. Der Kreis K, der durch Q geht und den Mittelpunkt M hat, enthält offenbar das gemeinsame Kreiszweieck von $K(M_1)$ und $K(M_2)$. Der von K und k begrenzte Kreisring enthält also die Kurve C. Sein Inhalt ist offenbar das π-fache der Potenz von Q in bezug auf k. Er ist also kleiner als der Inhalt des kleinsten Kreisringes von C vom Mittelpunkt M_1 oder M_2.

Aus diesem Widerspruch folgt, daß einer der Kreise $k(M_1)$ und $k(M_2)$ den anderen von innen berührt. Wir nehmen an, daß $k(M_2)$ den Kreis $k(M_1)$ in A von innen berührt.

$k(M_2)$ hat mindestens einen Punkt mit C gemeinsam. Daraus folgt, daß A auf C liegt, weil jeder andere Punkt von $k(M_2)$ im Innern von $k(M_1)$ und deshalb auch im Innern von B liegt.

Die Tangente von $k(M_1)$ in A ist eine Stützgerade von C und zugleich die Potenzlinie von $k(M_1)$ und $k(M_2)$. Bezeichnen wir mit $S(M_1)$ bzw. $S(M_2)$ den Kreisabschnitt, der von der Geraden QQ' aus $K(M_1)$ bzw. $K(M_2)$ abgeschnitten wird und B enthält, so liegt $S(M_2)$ offenbar in $S(M_1)$.

$K(M_1)$ hat mit $S(M_2)$ nur die Punkte Q und Q' und mit C mindestens einen Punkt gemeinsam. Daraus folgt, daß C mindestens einen der Punkte Q und Q' enthält.

Liegt nur Q auf C, so enthält der Rand von $K(M_1)$ keinen anderen Punkt von C und B. Es gibt also auf der Strecke QM_1 einen zu M_1 genügend nahe liegenden Punkt M', so daß B auch in dem durch Q gehenden Kreise K' vom Mittelpunkt M' enthalten ist. Bezeichnet ε den Abstand M_1M', so liegt der Kreis k' mit dem Mittelpunkt M' und Halbmesser $r(M_1)-\varepsilon$ in $k(M_1)$ und somit auch in B. Der von K' und k' begrenzte Kreisring enthält also C und hat den Inhalt $\pi\{[R(M_1)-\varepsilon]^2-[r(M_1)-\varepsilon]^2\}=\pi[R^2(M_1)-r^2(M_1)]-$ $-2\pi\varepsilon[R(M_1)-r(M_1)]$. Er hat also einen kleineren Inhalt, als der kleinste Kreisring von C mit dem Mittelpunkt M_1.

Aus diesem Widerspruch folgt, daß auch Q' auf C liegt.

Die Gerade QQ' ist eine Stützgerade von C, die Strecke QQ' gehört also zu C. Die Punkte Q und Q' sind Ecken von C, weil die Stützgeraden von $S(M_2)$ in Q und Q' auch C stützen.

Für den Beweis des ausgesprochenen Satzes bleibt nur der Fall zu erledigen, in dem $K(M_1)$ von $K(M_2)$ in Q berührt wird. Q liegt auf der Geraden M_1M_2 und hat bezüglich $k(M_1)$ und $k(M_2)$ dieselbe Potenz, weil die kleinsten Kreisringe von C mit den Mittelpunkten M_1 und M_2 gleichen Inhalt besitzen. Daraus folgt, daß Q mit dem Berührungspunkte A von $k(M_1)$ und $k(M_2)$ zusammenfällt.

Dies ist aber unmöglich, weil die konzentrischen Kreise $K(M_1)$ und $k(M_1)$ einander nicht berühren können.

Damit ist der ausgesprochene Satz vollständig bewiesen.

4. Ist M ein beliebiger Punkt der Strecke M_1M_2, so ist $R(M)=\overline{QM}$ und $r(M)=\overline{AM}$, weil $K(M)$ offenbar durch Q geht und $k(M)$ die Gerade QQ' berührt. Die Kreise $k(M_1)$, $k(M_2)$ und $k(M)$ gehören zu einem Kreisbüschel, sie haben also in bezug auf Q gleiche Potenz. Daraus folgt, daß $K(M)$ und $k(M)$ einen kleinsten Kreisring begrenzen.

Sind also M_1 und M_2 die äußersten Punkte der Geraden M_1M_2, die für C Mittelpunkte kleinster Kreisringe sind, so ist jeder Punkt der Strecke M_1M_2 der Mittelpunkt eines kleinsten Kreisringes von C.

Es gibt keinen kleinsten Kreisring von C, dessen Mittelpunkt außerhalb der Strecke M_1M_2 liegt.

Ist nämlich M_3 der Mittelpunkt eines kleinsten Kreisringes von C, so müssen die Kreise $k(M_1)$ und $k(M_2)$ von $k(M_3)$ im Punkte A berührt werden. Diese drei Kreise müssen auf derselben Seite der gemeinsamen Tangente QQ' liegen. Im entgegengesetzten Falle gäbe es nämlich unter $k(M_1)$, $k(M_2)$ und $k(M_3)$ ein Paar, von dem kein Kreis den anderen von innen berührt.

Daraus folgt der Satz:

Gehören mehrere kleinste Kreisringe zu einer konvexen Kurve C, so erfüllen ihre Mittelpunkte eine Strecke. Die äußeren Kreise der kleinsten Kreisringe von C haben eine gemeinsame Sehne s, die zu C gehört. Die inneren Kreise berühren die Kurve im Mittelpunkte von s.

5. Man kann leicht konvexe Kurven angeben, zu denen mehrere kleinste Kreisringe gehören.

Ist C die Begrenzung eines Kreisabschnittes S von einem Kreise K, so erfüllen die Mittelpunkte der kleinsten Kreisringe von C eine Strecke m der Mittellinie von S. Die Mittellinie von S bedeutet hier die Verbindungsstrecke der Halbierungspunkte des Kreisbogens und der Grundlinie von S. Ist S nicht kleiner als die Hälfte der Kreisfläche K, so ist m die Verbindungsstrecke des Mittelpunktes von K und des Halbierungspunktes der Mittellinie von S. Ist S kleiner als die Hälfte der Kreisfläche K, so liegen die Endpunkte vom m in den Halbierungspunkten der Grundlinie und der Mittellinie von S.

Ist C die Begrenzung eines gleichschenkligen Dreiecks, so liegen die Mittelpunkte der kleinsten Kreisringe — der Symmetrie wegen — auf der Höhe des Dreiecks. Ist die Grundlinie des Dreiecks größer als seine Schenkel, so erfüllen die Mittelpunkte der kleinsten Kreisringe von C die Verbindungsstrecke der Mittelpunkte des Hüllkreises (Pferchkreises) und des Inkreises des Dreiecks. Im entgegengesetzten Falle hat das gleichschenklige Dreieck nur einen kleinsten Kreisring.

6. *Der Mittelpunkt eines kleinsten Kreisringes der konvexen Kurve C kann nur dann auf C fallen, wenn mehrere kleinste Kreisringe zu C gehören.*

Liegt nämlich der Mittelpunkt M_0 eines kleinsten Kreisringes auf C, so fällt er mit dem Mittelpunkte M^* des Hüllkreises von C zusammen. Der entgegengesetzte Fall führt nämlich zu einem Widerspruch, da

$$(1)\quad \operatorname{Min} F(M) = R^2(M_0) - r^2(M_0) = R^2(M_0) > \\ > \operatorname{Min} R^2(M) = R^2(M^*) \geqq R^2(M^*) - r^2(M^*) = F(M^*).$$

Daraus folgt, daß $K(M_0)$ der Hüllkreis von C ist. Die Stützgerade von C in M_0 enthält einen Durchmesser QQ' von $K(M_0)$. Die Strecke QQ' gehört zu C. Im entgegengesetzten Falle wäre nämlich der kleinste Bogen des Hüllkreises $K(M_0)$, der jeden gemeinsamen Punkt mit C enthält, kleiner als ein Halbkreis von $K(M_0)$.[2])

Ist M_1 ein auf der Mittelgeraden des Punktpaares QQ' zu M_0 genügend nahe liegender Punkt von B, so ist offenbar

$$(2)\quad F(M_0) = R^2(M_0) = R^2(M_1) - r^2(M_1) = F(M_1).$$

Damit ist unsere Behauptung bewiesen[3]).

---

[2]) Bonnesen—Fenchel, S. 54.

[3]) Der entsprechende Satz für einen Minimalkreisring lautet folgendermaßen:

Ein Punkt der konvexen Kurve C ist kein Mittelpunkt des Minimalkreisringes.

Dieser Satz läßt sich ebenso beweisen, wie der obige Satz für kleinste Kreisringe.

Der Beweis führt hier statt (1) wegen

7. Sind M_1 und M_2 die Mittelpunkte von zwei kleinsten Kreisringen von C, so erhält man aus $F(M_1)=F(M_2)$ die Gleichung

$$R^2(M_2)-R^2(M_1)=r^2(M_2)-r^2(M_1).$$

Ist also $r(M_1)<r(M_2)$, so ist auch $R(M_1)<R(M_2)$.

Aus der Indentität

$$[R(M)-r(M)]\,.\,[R(M)+r(M)]=F(M)$$

folgt also, daß die Breite desjenigen kleinsten Kreisringes möglichst klein ist, der den größten inneren (und zugleich äußeren) Kreis besitzt.

Unter den kleinsten Kreisringen gibt es nur einen, dessen Breite möglichst klein ist, weil die inneren Kreise der kleinsten Kreisringe eine Gerade in einem Punkte berühren und auf derselben Seite dieser Geraden liegen. (Der innere Kreis als Punktkreis kann auch auf diese Gerade fallen.)

8. Nennt man *innere* bzw. *äußere Stützpunkte* von C die gemeinsamen Punkte von C mit dem inneren bzw. äußeren Kreise des kleinsten Kreisringes von minimaler Breite, so besteht der Satz:

Es gibt auf C keinen Bogen, der jeden äußeren Stützpunkt und keinen inneren enthält.

Aus einem Beweis in **3** (S. 177, fünfter Absatz) folgt, daß es auf C mindestens zwei äußere Stützpunkte gibt.

Für den Beweis nehmen wir an, daß C von den äußeren Stützpunkten P und P' in solche zwei Bogen C_1 und C_2 geteilt wird, von denen C_1 keinen inneren und C_2 keinen äußeren Stützpunkt enthält. Bezeichnet M_0 den Mittelpunkt des kleinsten Kreisringes von minimaler Breite, so teilen P und P' den Kreis $K(M_0)$ in zwei Bogen Γ_1 und Γ_2, von denen Γ_2 im Innern keinen (äußeren und inneren) Stützpunkt von C enthält. Die Zentralprojektion γ_1 bzw. γ_2 des Kreisbogens Γ_1 bzw. Γ_2 von M_0 aus auf $k(M_0)$ enthält keinen inneren Stützpunkt bzw. jeden inneren Stützpunkt von C.

$$\mathrm{Min}[R(M)-r(M)]=R(M_0)>\mathrm{Min}\,R(M)=R(M^*)\geqq R(M^*)-r(M^*)$$

zu einem Widerspruch.

Statt (2) hat man jetzt die Ungleichung

$$R(M_0)-r(M_0)=R(M_0)>R(M_1)-r(M_1),$$

weil $R(M_0)$, $R(M_1)$ und $r(M_1)$ die Seiten des Dreiecks M_0M_1Q sind.

Beim Beweis haben wir hier die Heranziehung unendlich kleiner Größen von höherer Ordnung vermieden, die beim Beweis von Kritikos (S. 584) eine wesentliche Rolle spielen.

Γ_1 und ein beliebiger innerer Stützpunkt können nicht auf derselben Seite der Geraden PP' liegen. Sonst fiele γ_2 in das offenbar zu B gehörige Dreieck M_0PP'. Dann könnte aber γ_2 und damit $k(M_0)$ keinen Stützpunkt von C enthalten.

Bezeichnet T bzw. T' den Berührungspunkt der von P bzw. P' an γ_2 gehenden Tangente, so liegen die inneren Stützpunkte von C auf dem zwischen T und T' liegenden abgeschlossenen Teilbogen δ_2 von γ_2. Die inneren Punkte der beiden anderen Teilbogen von γ_2 liegen nämlich im Innern der konvexen Hülle des Punktpaares PP' und des Kreises $k(M_0)$ und deshalb auch im Innern von B.

Der Punkt T ist also nur dann ein Stützpunkt, wenn die Strecke PT ein Teil von C ist.

Fallen T und T' zusammen, so gehört die Strecke PP' der Kurve C an. Dann berührt $k(M_0)$ die Strecke PP' in T und hat mit C keinen anderen Punkt gemeinsam. Offenbar gibt es einen zu M_0 hinreichend nahe liegenden Punkt M_1, so daß $k(M_1)$ in B liegt, den Kreis $k(M_0)$ enthält und in T berührt. Nach **4** begrenzen $K(M_1)$ und $k(M_1)$ einen kleinsten Kreisring, der nach **7** eine kleinere Breite hat, als der kleinste Kreisring von minimaler Breite (mit dem Mittelpunkt M_0). Aus diesem Widerspruch folgt die Richtigkeit des obigen Satzes, falls T und T' zusammenfallen.

Ist T kein Stützpunkt, so läßt sich P durch einen hinreichend nahe liegenden Punkt P_1 von Γ_2 so ersetzen, daß der Berührungspunkt T_1 auf der von P_1 an δ_2 gehenden Tangente kein Stützpunkt ist und daß es auf δ_2 zwischen T und T_1 keinen Stützpunkt gibt. Ist T ein Stützpunkt, so fällt P_1 bzw. T_1 mit P bzw. T zusammen. Auf ähnliche Weise verfährt man, falls T' kein Stützpunkt ist.

Nimmt man die Punkte P_1, P_1', T_1 und T_1' auf diese Weise an, so enthält B den zwischen T und T' fallenden Teilbogen δ_2' von δ_2 und je eine Teilstrecke von P_1T_1 und $P_1'T_1'$. Der Kreis $k(M_0)$ hat mit C außerhalb von δ_2' keinen gemeinsamen Punkt.

P_1 (oder P_1') ist entweder kein Punkt von C oder eine Ecke von C. Daraus folgt, daß man durch P_1 und P_1' einen zu $K(M_0)$ hinreichend nahen größeren Kreis K_1 mit dem Mittelpunkte M_1 führen kann, der B enthält. Offenbar kann man M_1 auch so wählen, daß derjenige Kreis k_1 mit dem Mittelpunkte M_1, der die Geraden P_1T_1 und $P_1'T_1'$ in den Punkten P_2 und P_2' berührt, im Bereich B

liegt. Der von K_1 und k_1 begrenzte Kreisring enthält C und hat nach **3** den Inhalt $\pi\overline{P_1T_2}^2$. Dieser Inhalt ist also kleiner, als der Inhalt $\pi\overline{P_1T_1}^2$ des kleinsten Kreisringes mit dem Mittelpunkt M_0.

Aus diesem Widerspruch folgt der obige Satz[4]).

9. Der Zusammenhang zwischen dem Minimalkreisring und den kleinsten Kreisringen von C wird durch den folgenden Satz beleuchtet:

Hat eine konvexe Kurve C nur einen kleinsten Kreisring, so ist er ihr Minimalkreisring. Hat aber C mehrere kleinste Kreisringe, so stimmt ihr Minimalkreisring mit demjenigen kleinsten Kreisring überein, dessen innerer Kreis am größten ist.

Dieser Satz folgt aus dem folgenden Satze von T. BONNESEN und N. KRITIKOS[5]):

Liegt die konvexe Kurve C in dem von den konzentrischen Kreisen $\overline{K}$ und $\overline{k}$ begrenzten Kreisring vom Mittelpunkt $\overline{M}$ und lassen sich die gemeinsamen Punkte der Kurven C und $\overline{K}$ von den Zentralprojektionen der gemeinsamen Punkte von C und $\overline{k}$ auf $\overline{K}$ von $\overline{M}$ aus durch keine Gerade der Ebene trennen, so ist der Kreisring von $\overline{K}$ und $\overline{k}$ eindeutig bestimmt. Er fällt mit dem Minimalkreisring von C zusammen.

Aus dem Satze für die Lage der Stützpunkte von C mit dem kleinsten Kreisringe von minimaler Breite folgt einfach, daß dieser

---

[4]) Dieser Satz gilt ohne Weiteres auch für die Lage der Stützpunkte von C mit den zwei Kreisen des Minimalkreisringes.

Gilt nämlich der Satz für den Minimalkreisring nicht, so wird C von zwei Punkten P_1 und P_1' in solche Bogen C_1 und C_2 geteilt, von denen C_1 keinen inneren und C_2 keinen äußeren Stützpunkt von C enthält. Wir bezeichnen mit M_0 den Mittelpunkt des Minimalkreisringes, mit Γ_1 und Γ_2 bzw. γ_1 und γ_2 die Zentralprojektionen von C_1 und C_2 von M_0 aus auf $K(M_0)$ bzw. $k(M_0)$, mit P bzw. p die Projektion von P_1 auf $K(M_0)$ bzw. $k(M_0)$ in derselben Zentralprojektion und mit D den Halbierungspunkt von Γ_1.

Ist nun M_1 ein zu M_0 genügend nahe liegender Punkt der Strecke M_0D, so hat man offenbar $R(M_1) \leqq \overline{PM_1}$ und $r(M_1) \geqq \overline{pM_1}$. Dies führt aber zum Widerspruch

$$\mathrm{Min}\,[R(M)-r(M)] = R(M_0)-r(M_0) = \overline{Pp} > \overline{PM_1} - \overline{pM_1} \geqq R(M_1) - r(M_1),$$

Aus diesem Widerspruch folgt der Satz für den Minimalkreisring von C.

Dieser Satz und ein entsprechender von KRITIKOS (S. 584) lassen sich aus einander ableiten. Beim Beweis verwendet aber KRITIKOS unendlich kleine Größen höherer Ordnung.

[5]) KRITIKOS, S. 586.

Kreisring die voranstehenden Eigenschaften des Kreisringes von $\bar{K}$ und $\bar{k}$ besitzt und deshalb mit dem Minimalkreisring von C übereinstimmt.

10. Die für konvexe Kurven erhaltenen Resultate lassen sich auch auf konvexe Flächen verallgemeinern.

Als *Kugelschale* wird die abgeschlossene Menge zwischen zwei konzentrischen Kugeloberflächen bezeichnet. Eine Kugelschale, die eine konvexe Fläche F enthält und für welche die Differenz der Flächeninhalte ihrer zwei Kugeln bzw. die Differenz der Halbmesser ihrer zwei Kugeln am kleinsten ist, soll als *Kugelschale von kleinster Flächenabweichung* bzw. als *Minimalkugelschale* bezeichnet werden, wenn die innere Kugel innerhalb der Fläche F liegt[6]).

Die Existenz wenigstens einer Kugelschale von kleinster Flächenabweichung zu einer konvexen Fläche erkennt man ebenso, wie die Existenz eines kleinsten Kreisringes zu einer konvexen Kurve.

Auf Grund des Obigen kann man ohne größere Schwierigkeit den folgenden Satz beweisen:

Hat eine konvexe Fläche F mehrere Kugelschalen von kleinster Flächenabweichung, so berühren sich die inneren Kugeln dieser Kugelschalen in einem Punkte M von F. Die äußeren Kugeln gehen durch einen Kreis K vom Mittelpunkt M, der in der zu M gehörigen Stützebene von F liegt. In dieser Stützebene enthält F entweder einen Durchmesser von K, oder einen konvexen Bereich mit dem Hüllkreise K. Die Minimalkugelschale ist von kleinster Flächenabweichung. Hat F mehrere Kugelschalen von kleinster Flächenabweichung, so stimmt die Minimalkugelschale von F mit derjenigen Kugelschale von kleinster Flächenabweichung überein, deren innree Kugel am größten ist[7]).

Ist F die Begrenzung eines Kugelabschnittes, dessen Grund-

---

[6]) Bonnesen—Fenchel, S. 54.

[7]) Ich möchte hier bemerken, daß einige Beweise für Minimalkugelschalen bei Kritikos sich vereinfachen lassen.

Der Beweisgang in der Fußnote [3]) führt auch zum Satz: Ein Punkt von F ist kein Mittelpunkt der Minimalkugelschale.

Nennt man innere bzw. äußere Stützpunkte der konvexen Fläche F ihre gemeinsamen Punkte mit der inneren bzw. äußeren Kugeloberfläche der Minimalkugelschale, so gilt der Satz von Kritikos:

Projiziert man vom Mittelpunkt M_0 der Minimalkugelschale von F aus kreis K den Mittelpunkt M hat, so hat F offenbar mehrere Kugelschalen von kleinster Flächenabweichung. Die inneren Kugeloberflächen berühren sich in M, die äußeren gehen durch K.

Ist K der Durchschnittskörper eines Drehzylinders Z und einer Kugel S vom gleichen Durchmesser, deren Mittelpunkt M auf dem Mantel von Z liegt, so ist die Begrenzung von K eine konvexe Fläche F, die offenbar mehrere Kugelschalen von kleinster Flächenabweichung besitzt. Die inneren Kugeln dieser Kugelschalen berühren den Zylindermantel in M, die äußeren Kugeln gehen durch den Hauptkreis von S, der in der zu M gehörigen Berührungsebene des Zylindermantels liegt. Diese Fläche F enthält nur eine Strecke, einen Durchmesser von S.

(Eingegangen am 19. Januar 1942.)

---

die inneren Stützpunkte auf die äußere Kugeloberfläche $O(M_0)$, so sind diese Projektionen durch keine Ebene von den äußeren Stützpunkten getrennt.

Der Beweisgang in der Fußnote [4]) läßt sich auch zum Beweis dieses Satzes anwenden.

Besteht nämlich dieser Satz nicht zurecht, so gibt es eine Ebene ε, von der $O(M_0)$ in zwei Kugelkappen O_1 und O_2 geteilt wird, so daß das Innere von O_1 bzw. O_2 jeden äußeren Stützpunkt bzw. die Projektion jedes inneren Stützpunktes enthält. Wir bezeichnen mit P bzw. p einen Schnittpunkt von $O(M_0)$ mit ε bzw. seine Zentralprojektion auf die innere Kugeloberfläche $o(M_0)$ und mit D den von ε am weitestens liegenden Punkt von O_1. Dann gibt es auf der Strecke M_0D offenbar einen zu M_0 genügend nahe liegenden Punkt M_1, so daß F in der durch P gehenden Kugel S_1 vom Mittelpunkt M_1 und außerhalb der durch p gehenden Kugel s_1 von demselben Mittelpunkt liegt.

Die Differenz der Halbmesser von S_1 und s_1 ist aber kleiner, als diejenige von $O(M_0)$ und $o(M_0)$. Aus diesem Widerspruch folgt der Satz.

Dieser Satz von Kritikos gilt auch für diejenige Kugelschale von minimaler Flächenabweichung, die die größte innere Kugel besitzt. Der Beweis geschieht durch eine entsprechende Abänderung des für den kleinsten Kreisring von minimaler Breite gegebenen Beweises.

On a theorem of S. Bernstein.

By G. GRÜNWALD (†) in Ujpest.

Let $f(x)$ be continuous in the interval $-1 \leqq x \leqq +1$ and let us denote

(1) $$L_n[f; x] \qquad n=1, 2, \ldots$$

the n-th Lagrange interpolation polynomial of the function $f(x)$ corresponding to the Chebysheff abscissas. Then (1) is the unique polynomial of degree $n-1$ which assumes the values $f(x_1^{(n)})$, $f(x_2^{(n)}), \ldots, f(x_n^{(n)})$ at the zeros

(2) $$x_1^{(n)}=\cos\frac{\pi}{2n},\ x_2^{(n)}=\cos 3\frac{\pi}{2n}, \ldots, x_n^{(n)}=\cos(2n-1)\frac{\pi}{2n}$$

of the n-th Chebysheff polynomial $T_n(x)=\cos(n \arccos x)$. It is known that the sequence (1) is not convergent for every continuous function[1]); moreover there exist continuous functions for which the sequence (1) is divergent everywhere[2]) in the interval $-1 \leqq x \leqq +1$. S. BERNSTEIN proved[3]) that we can modify (1) by choosing suitably λn of the abscissas (2) and changing there suitably the values of f, so that the new sequence of interpolation polynomials be

---

[1]) G. FABER, Über die interpolatorische Darstellung stetiger Funktionen, *Jahresbericht der Deutschen Math. Vereinigung*, **23** (1914), p. 190–210; S. BERNSTEIN, Sur la limitation des valeurs ..., *Bulletin Academy of Sciences de l'URSS*, 1931, p. 1025–1050.

[2]) G. GRÜNWALD, Über Divergenzerscheinungen der Lagrangeschen Interpolationspolynome stetiger Funktionen, *Annals of Math.*, **37** (1936), p. 908–918. See also J. MARCINKIEWICZ, Sur la divergence des polynomes d'interpolation, *these Acta*, **7** (1937), p. 131–135.

[3]) S. BERNSTEIN, Sur une modification de la formule d'interpolation de Lagrange, *Communications de la Société Math. de Kharkow*, (4) **5** (1932), p. 49–57.

uniformly convergent in the interval $-1 \leqq x \leqq +1$ for every continuous function, where $\lambda > 0$ is an arbitrary small but fixed number.

In this note, we prove that the theorem of BERNSTEIN is exact in the sense that, instead of fixed λ, we cannot take a sequence $\lambda = \lambda_n$ for which $\lambda_n \to 0$ if $n \to \infty$. In other words, if $\lambda_n \to 0$ for $n \to \infty$, if we choose $n\lambda_n$ abscissas from (2) and we change the values of f assumed here arbitrarily, then there exist continuous functions for which the modified sequence of interpolation polynomials is not uniformly convergent in the interval $-1 \leqq x \leqq +1$.

The following explicit form of (1) is known:

$$(3) \quad L_n[f; x] = \sum_{k=1}^{n} (-1)^{k+1} \frac{T_n(x)\sqrt{1-(x_k^{(n)})^2}}{n(x-x_k^{(n)})} f(x_k^{(n)}) = \sum_{k=1}^{n} l_k(x) f(x_k^{(n)}).$$

The modified interpolation polynomial is

$$(4) \qquad B_n[f; x] = \sum_{k=1}^{n} A_k^{(n)} l_k(x),$$

where

$$(5) \qquad A_{i_1}^{(n)}, A_{i_2}^{(n)}, \ldots, A_{i_\mu}^{(n)} \qquad \mu = n\lambda_n$$

are arbitrary numbers and

$$(6) \qquad A_{i_{\mu+1}}^{(n)} = f(x_{i_{\mu+1}}^{(n)}),\ A_{i_{\mu+2}}^{(n)} = f(x_{i_{\mu+2}}^{(n)}), \ldots, A_{i_n}^{(n)} = f(x_{i_n}^{(n)})$$

($i_1, i_2, \ldots, i_n$ being an arrangement of the numbers $1, 2, \ldots, n$).

Evidently we can suppose that

$$(7) \qquad \overline{\lim}\, |A_k^{(n)}| \leqq 1, \qquad k = 1, 2, \ldots, \mu = \lambda_n n;\ n = 1, 2, \ldots$$

Let us divide the interval $-1 \leqq x \leqq +1$ into $\mu + 1$ equal parts. There is at least one part not containing any of the abscissas $x_{i_1}^{(n)}, x_{i_2}^{(n)}, \ldots, x_{i_\mu}^{(n)}$. Let $x_j^{(n)}, x_{j+1}^{(n)}$ be the two Tschebyscheff abscissas intercepting the centre of this interval and put $x_0 = \cos j\frac{\pi}{n}$. Then we have[4])

$$(8) \quad \sum_{k=1}^{\mu} |l_{i_k}(x_0)| = \frac{1}{n} \sum_{k=1}^{\mu} \frac{\sqrt{1-(x_{i_k}^{(n)})^2}}{|x_0 - x_{i_k}^{(n)}|} <$$

$$< \frac{1}{n} \sum_{k=1}^{\mu} \frac{1}{|x_0 - x_{i_k}|} < \frac{c}{n} \sum_{k=1}^{\mu} \frac{n}{\mu + k} < c.$$

---

[4]) $c > 0$ denotes an absolute constant, not necessarily the same one each time it occurs.

Let $f_n(x)$ be a continuous function in the interval $-1 \leqq x \leqq +1$ for which

(9) $$f_n(x_k^{(n)}) = \operatorname{sign} l_k(x_0), \qquad k = 1, 2, \ldots, n.$$

From (4), (5), (6), (7), (8), (9) it follows that

$$|B_n[f_n; x_0]| = \left| \sum_{k=1}^{n} A_k^{(n)} l_k(x_0) \right| =$$

(10) $$= \left| \sum_{k=1}^{n} f_n(x_k^{(n)}) l_k(x_0) + \sum_{k=1}^{\mu} (A_{i_k}^{(n)} - f(x_{i_k}^{(n)}))\, l_{i_k}(x_0) \right| \leqq$$

$$\leqq \sum_{k=1}^{n} |l_k(x_0)| - 2 \sum_{k=1}^{\mu} |l_{i_k}(x_0)| < \sum_{k=1}^{n} |l_k(x_0)| - c.$$

Since[5])

(11) $$\sum_{k=1}^{n} |l_k(x)| > \frac{1}{\pi} \log n\, |T_n(x)|$$

and so

(12) $$\sum_{k=1}^{n} |l_k(x_0)| > \frac{1}{\pi} \log n,$$

we have

(13) $$|B_n[f_n; x_0]| > c \log n.$$

With the aid of (13) it is easy to prove that the sequence of interpolation polynomials $B_n[f; x]$ is not uniformly convergent if $f(x)$ is the following continuous function

(14) $$f(x) = \sum_{\nu=1}^{\infty} \frac{f_{n_\nu}(x)}{\nu^2}$$

and if the sequence n_ν is suitably chosen.

(Received September 8, 1940.)

---

[5]) G. Grünwald, Über Divergenzerscheinungen der Lagrangeschen Interpolationspolynome, *these Acta*, 7 (1935), p. 206–221.

Über die Wurzeln der Dirichletschen L-Funktionen.

Von P. TURÁN in Budapest.

Im folgenden sei k stets eine positive ganze Zahl, $\varphi(k)$ die Eulersche zahlentheoretische Funktion, $\chi(n)$ eine beliebige von den zu k gehörigen Dirichletschen Charakteren (deren Anzahl bekanntlich $\varphi(k)$ ist). Die komplexe Veränderliche sei $s=\sigma+it$, α ein reeller Parameter und $a_1, a_2, \ldots$ Konstanten, welche von s, k und α unabhängig sind. Die Dirichletschen Funktionen $L(s,\chi)$ sind bekanntlich diejenige analytischen Funktionen, welche für $\sigma>1$ durch die Reihe $\sum_{n=1}^{\infty}\frac{\chi(n)}{n^s}$ definiert sind.

Schon DIRICHLET erkannte im Jahre 1837 den Zusammenhang zwischen den Wurzeln der L-Funktionen und der Anzahl der Primzahlen der arithmetischen Progression $kx+l$, wo $(k,l)=1$; der schwerste Hilfssatz seines berühmten Satzes über die Existenz unendlich vieler Primzahlen der Progression war eben der Nachweis, daß $L(1,\chi)\neq 0$ für jedes χ und k. Der Dirichletsche Satz wurde sechzig Jahre später von DE LA VALLEE – POUSSIN in eine asymptotische Formel für die Primzahlen der Progression $kx+l$ verfeinert; dazu benötigte er aber die schärfere Tatsache, daß $L(s,\chi)\neq 0$ auf der ganzen Gerade $\sigma=1$ gilt. Diese Tatsache können auch die neueren Methoden von LANDAU und N. WIENER nicht entbehren. Wenn man zu der asymptotischen Primzahlformel auch ein Restglied zufügen will, muß man auch die komplexen Nullstellen der L-Funktionen im „kritischen" Streifen $0<\sigma<1$ betrachten. Bekanntlich hat hier jede der L-Funktionen unendlich viele Wurzeln; es ist sehr wahrscheinlich, daß alle diese auf der

Geraden $\sigma = \frac{1}{2}$ liegen[1]). Dies ist nicht einmal für den einfachsten Fall bewiesen, wenn $\chi = \chi_0$, der Hauptcharakter ist; in diesem Falle reduziert sich die Frage nach der Formel

$$(1) \qquad L(s, \chi_0) = \zeta(s) \prod_{p/k} \left(1 - \frac{1}{p^s}\right),$$

(wo $\zeta(s)$ die Riemannsche Funktion bedeutet) auf die klassische Riemannsche Vermutung $\zeta(s) \neq 0$ für $\sigma > \frac{1}{2}$. Das weitgehendste Resultat in dieser Richtung verdanken wir TITCHMARSH[2]), nach welchem ein a_1 existiert so, daß im Gebiete $\sigma \geqq 1 - a_1 \log^{-0.81}(|t| + 2)$ $\zeta(s) \neq 0$ gilt.

Die meisten Ergebnisse der Theorie der L-Funktionen beziehen sich bei festem k auf $|t| \to \infty$. Die Klassenanzahl der definiten binären quadratischen Formen mit der Discriminante $-k$, die obere Abschätzung der kleinsten Primzahl der Progression $kx + l$ führten zu der Frage, wie sich die L-Funktionen im Gebiete $\frac{1}{2} \leqq \sigma \leqq 1$, $|t| \leqq 5$ für $k \to \infty$ benehmen. Aus den vielen schönen Untersuchungen in dieser Richtung heben wir nur die folgenden hervor. Ein Satz von LITTLEWOOD[3]) wirft Licht auf die wichtige Frage, ob die L-Funktionen für $\sigma \geqq \frac{1}{2}$ reelle Wurzeln besitzen oder nicht; er besagt, daß jede L-Funktion mod k wenigstens eine Wurzel im Gebiete $\sigma \geqq \frac{1}{2}$, $|t| \leqq \frac{1}{\log \log \log k}$ besitzt. Aus einem Satz von SIEGEL[4]) — wie WALFISZ[5]) bemerkte — folgt, daß es eine, nur von ε abhängige Konstante $A(\varepsilon)$ gibt so, daß keine der zu k gehörigen L-Funktionen auf der Strecke $1 \geqq \sigma \geqq 1 - A(\varepsilon) k^{-\varepsilon}$

1) Explizit ausgesprochen bei A. PILTZ, *Habilitationsschrift*, 1884. Siehe auch HARDY and LITTLEWOOD, Some Problems of Partitio Numerorum III, *Acta Math.*, **44** (1922), p. 1—70.

2) E. C. TITCHMARSH, On $\zeta(s)$ and $\pi(x)$, *Quarterly Journal of Math.*, Oxford Series, **9** (1938), p. 97—108.

3) J. E. LITTLEWOOD, Researches in the theory of the Riemann ζ-function, *Proceedings London Math. Society*, (2) **20** (1922), Records, p. XXII—XXVIII.

4) C. L. SIEGEL, Über die Classenzahl quadratischer Zahlkörper, *Acta Arithmetica*, **1** (1936), p. 83—86.

5) A. WALFISZ, Zur additiven Zahlentheorie II, *Math. Zeitschrift*, **40** (1936), p. 592—607.

reelle Wurzeln besitzt. In Verschärfung eines Gronwallschen[6]) Resultats bewies A. PAGE[7]) schon früher, daß es ein a_2 gibt so, daß auf der Strecke $1 \geqq \sigma \geqq 1 - a_2 \log^{-1} k$ höchstens eine der L-Funktionen $\bmod k$ verschwinden kann. Diese gehört dann zu einem reellen Charakter und besitzt hier eine einzige einfache Wurzel. Die von PAGE benützte Methode stammt von LANDAU[8]); er und anschließend TITCHMARCH[9]) bewiesen, daß für ein geeignetes a_3 im Gebiete $\sigma \geqq 1 - a_3 \log^{-1} k$, $|t| \leqq 5$ keine der zu komplexen Charakteren gehörigen L-Funktionen verschwindet, ferner, daß die zu reellen Charakteren gehörigen L-Funktionen hier nur auf der reellen Achse verschwinden können. Aber es ist bisher — abgesehen von $L(s, \chi_0)$ — für keine einzige L-Funktion $\bmod k$ bekannt, daß sie im Rechtecke $\sigma > \frac{1}{2}$, $|t| \leqq 5$ (oder auch nur in einem Rechtecke $\sigma \geqq 1 - a_4$, $|t| \leqq 5$) nicht verschwindet. (Was $L(s, \chi_0)$ betrifft, sie besitzt nach (1) überhaupt für $\sigma > 0$, $|t| \leqq 14$ keine Wurzeln, da die Wurzel von $\zeta(s)$ mit dem kleinsten positiven imaginären Teil bekanntlich[10]) $\frac{1}{2} + i \,.\, 14{,}13 \ldots$ ist.)

Im folgenden beweise ich, daß solche L-Funktionen $\bmod k$ wirklich existieren und sogar *die Anzahl derjenigen L-Funktionen* $\bmod k$, *welche im Gebiete* $\sigma \geqq \frac{1}{2} + \frac{5}{\sqrt{\log \log k}}$, $|t| \leqq 5$ *wenigstens eine Wurzel haben, durch die Gesamtanzahl der L-Funktionen* $\bmod k$, *d. h. durch* $\varphi(k)$ *dividiert, gegen* 0 *strebt.* Also ist die verallgemeinerte Riemannsche Vermutung „im Wesentlichen" für $|t| \leqq 5$ für „fast alle" L-Funktionen $\bmod k$ wahr. Sätze I und II werden noch schärfere Abschätzungen für die Anzahl derjenigen L-Funktionen $\bmod k$ ergeben, welche im Gebiete $\sigma \geqq \alpha$, $|t| \leqq 5$ wenigstens eine Wurzel besitzen und sogar für die Gesamtanzahl der hier

[6]) H. GRONWALL, Sur les séries de Dirichlet correspondant à des caractères complexes, *Rendiconti Palermo,* 35 (1913), p. 145–159.

[7]) A. PAGE, On the number of primes in an arithmetic progression, *Proceedings London Math. Society,* (2) 39 (1935), p. 116–141.

[8]) E. LANDAU, Über das Nichtverschwinden der Dirichletschen Reihen, welche komplexen Charakteren entsprechen, *Math. Annalen,* 70 (1911), p. 69–78.

[9]) E. C. TITCHMARCH, On a divisor problem, *Rendiconti Palermo,* 54 (1930), p. 414–429.

[10]) Siehe z. B. E. C. TITCHMARSH, *The zeta-function of Riemann* (Cambridge, 1930), p. 45.

liegenden Wurzeln aller L-Funktionen $\bmod k$; hier bedeutet α eine beliebige Zahl im Intervalle $\frac{1}{2}+\frac{5}{\sqrt{\log\log k}} \leqq \alpha \leqq 1-\frac{2}{\log k}$. Diese Sätze erinnern uns interessanterweise an einen Satz von CARLSON[11]), welche sich auf die Wurzelverteilung der Riemannschen ζ-Funktion für $|t| \to \infty$ bezieht.

In einem früheren Aufsatze[12]) bewies ich, daß, wenn wir die kleinste Primzahl der Progression $l, l+k, l+2k, \ldots$ mit $P(k, l)$ bezeichnen, dann $P(k, l) < a_4 \varphi(k)^{a_5}$ gilt, vorausgesetzt, daß ein $\frac{1}{2} \leqq a_6 < 1$ existiert mit der Eigenschaft, daß keine der L-Funktionen $\bmod k$ im Gebiete $\sigma \geqq a_6$, $|t| \leqq 5$ verschwindet. Hier hängt a_5 natürlich von a_6 ab. Neuerdings habe ich die Ungleichung $P(k, l) < a_4 \varphi(k)^{a_5}$ auf Grund der wesentlich schwächeren Annahme bewiesen, daß es ein a_7 gibt so, daß keine der L-Funktionen $\bmod k$ im Gebiete

$$\sigma \geqq 1 - a_7 \frac{\log\log k}{\log k} \qquad (2)$$
$$|t| \leqq 5$$

verschwindet. Satz I gibt Orientierung für die Wahrscheinlichkeit dieser Vermutung, da die Anzahl *aller* Wurzeln *aller* L-Funktionen $\bmod k$ im Gebiete $\sigma \geqq \frac{1}{4}$, $|t| \leqq 5$ gewiß $> a_8 \varphi(k) \log k$ und von diesen nach Satz I im Teilgebiet (2) nur höchstens $a_9 \log^{8a_7+6} k$ liegen Auf diese Frage werde ich in einer anderen Arbeit zurückkommen.

Im folgenden bezeichnen wir mit $N_k(\alpha)$ die Anzahl derjenigen Wurzeln *aller* L-Funktionen $\bmod k$, welche im Rechtecke $2 \geqq \sigma \geqq \alpha$, $|t| \leqq 5$ liegen. Es sei stets $k > e^{30}$. Dann beweisen wir die folgenden zwei Sätze.

Satz I. *Es sei* $1-\frac{2}{\log k} \geqq \alpha \geqq \frac{-1+\sqrt{8}}{2}$. *Dann gilt*

$$N_k(\alpha) < a_9 k^{4(1-\alpha^2)} \log^{\frac{11}{2}} k.$$

[11]) F. CARLSON, Über die Nullstellen der Dirichletschen Reihen und der Riemannschen ζ-Funktion, *Arkiv för Math., Astronomi och Fysik,* **15** (1920), No. 20.

[12]) P. TURÁN, Über die Primzahlen der arithmetischen Progression (II), *diese Acta,* **9** (1939), p. 187—192.

Satz II. *Es sei* $\frac{-1+\sqrt{8}}{2} \geqq \alpha \geqq \frac{1}{2} + \frac{5}{\sqrt{\log\log k}}$. *Dann gilt*

$$N_k(\alpha) < a_{10} k^{\frac{1}{2}+2\alpha-2\alpha^2} \log k.$$

Beide der Exponenten $4(1-\alpha^2)$ und $\frac{1}{2}+2\alpha-2\alpha^2$ sind für die betrachteten α-Werte kleiner als 1. Die Exponenten könnte man mit umständlicheren Rechnungen, nämlich mittels bessere Wahl von zwei Parametern, etwas verkleinern.

Zum Beweis benötigen wir einige Lemmata.

Lemma I.[13]) *Es sei* $1-\frac{2}{\log k} \geqq \alpha \geqq \frac{1}{2}$, $\sigma \geqq \alpha - \frac{1}{\log k}$ *und*

$$S_l = \sum_{\substack{n \equiv l \bmod k \\ x \leqq n \leqq y}} \frac{d(n)}{n^\sigma},$$

wo $d(n)$ *die Anzahl der Teiler von* n *bedeutet,* x *und* y *Abkürzungen für* $k^{2\left(\alpha-\frac{1}{\log k}\right)}$ *bzw.* $k^{2\left(1+\alpha-\frac{1}{\log k}\right)}$ *sind. Dann gilt*

$$S = \sum_{l=1}^{k} S_l^2 < a_{11} k^{4(1-\alpha^2)-1} \log^5 k.$$

Beweis: Für jedes l gilt offenbar[14])

$$S_l^2 = \left(\sum_{\substack{n\equiv l \\ x \leqq n \leqq y}} \frac{d(n)}{n^{\frac{\sigma}{2}}} \frac{1}{n^{\frac{\sigma}{2}}} \right)^2 \leqq \left(\sum_{\substack{n\equiv l \\ x \leqq n \leqq y}} \frac{d(n)^2}{n^\sigma} \right) \left(\sum_{\substack{n\equiv l \\ x \leqq n \leqq y}} \frac{1}{n^\sigma} \right). \tag{3}$$

Da

$$\sum_{\substack{n\equiv l \\ x \leqq n \leqq y}} \frac{1}{n^\sigma} \leqq \frac{1}{x^\sigma} + \frac{1}{k} \int_x^y \frac{du}{u^\sigma} \leqq a_{12} \left(k^{-2\alpha^2} + \frac{1}{k} \int_1^y \frac{du}{u^\alpha} \right) <$$
$$< a_{13} (k^{-2\alpha^2} + k^{2(1+\alpha)(1-\alpha)-1} \log k) < a_{14} k^{2(1-\alpha^2)-1} \log k \tag{4}$$

gilt, folgt aus (3) und (4)

$$S \leqq a_{14} k^{2(1-\alpha^2)-1} \log k \sum_{1 \leqq n \leqq y} \frac{d(n)^2}{n^\alpha}. \tag{5}$$

---

[13]) Nachträglich habe ich bemerkt, daß Lemma I aus Lemma II folgt und sogar mit $\log^4 k$ statt $\log^5 k$.

[14]) Wir verlassen von nun an die Bezeichnung „mod k“ überall wo nicht ein Mißverständnis zu befürchten ist.

Bekanntlich gilt

$$\sum_{n \leqq N} d(n)^2 = O(N \log^3 N),$$

also gewinnt man durch partielle Summation

$$(6) \qquad \sum_{n \leqq y} \frac{d(n)^2}{n^\alpha} \leqq a_{15} \int_1^y \frac{\log^3 u}{u^\alpha}\, du < a_{16} k^{2(1-\alpha^2)} \log^4 k.$$

(5) und (6) beweisen Lemma I.

Lemma II. *Wenn* $(k, l) = 1$ *und* $\frac{1}{2} \leqq \alpha \leqq 1 - \frac{2}{\log k}$, *dann gilt für* $\sigma \geqq \alpha - \frac{1}{\log k}$ *die Ungleichung*

$$S_l < a_{17} k^{1-2\alpha^2} \log^2 k.$$

Beweis. Es ist, mit den Bezeichnungen des vorigen Lemmas,

$$S_l = \sum_{\substack{x \leqq n \leqq y \\ n \equiv l}} \frac{1}{n^\sigma} \Big(\sum_{d/n} 1\Big) \leqq 2 \sum_{\substack{x \leqq n \leqq y \\ n \equiv l}} \frac{1}{n^\sigma} \Big(\sum_{\substack{d/n \\ d \leqq \sqrt{n}}} 1\Big) =$$

$$= 2 \sum_{\substack{d \leqq \sqrt{y} \\ (d, k) = 1}} \frac{1}{d^\sigma} \left(\sum_{\substack{\max\left(d, \frac{x}{d}\right) \leqq m \leqq \frac{y}{d} \\ dm \equiv l}} \frac{1}{m^\sigma} \right) =$$

$$= 2 \sum_{\substack{d \leqq \sqrt{x} \\ (d, k) = 1}} \frac{1}{d^\sigma} \left(\sum_{\substack{\frac{x}{d} \leqq m \leqq \frac{y}{d} \\ m \equiv l_1}} \frac{1}{m^\sigma} \right) +$$

$$+ 2 \sum_{\substack{\sqrt{x} < d \leqq \sqrt{y} \\ (d, k) = 1}} \frac{1}{d^\sigma} \left(\sum_{\substack{d \leqq m \leqq \frac{y}{d} \\ m \equiv l_1}} \frac{1}{m^\sigma} \right) = S_l' + S_l'',$$

wo l_1 die (nach $(d, k) = 1$ immer existierende) Lösung der Kongruenz $dl_1 \equiv l \bmod k$ bedeutet. Dann ist aber, wie in (4),

$$S_l' < a_{18} \sum_{d \leqq \sqrt{x}} \frac{1}{d^\alpha} \left(\frac{d^\alpha}{x^\alpha} + \frac{1}{k} \int_1^{\frac{y}{d}} \frac{du}{u^\alpha} \right) <$$

$$(7) \qquad < a_{19} \left(x^{\frac{1}{2} - \alpha} + \frac{y^{1-\alpha}}{k(1-\alpha)} \sum_{d \leqq \sqrt{x}} \frac{1}{d} \right) <$$

$$< a_{20} \left(k^{\alpha - 2\alpha^2} + k^{2(1+\alpha)(1-\alpha)-1} \log^2 k\right) < a_{21} k^{2(1-\alpha^2)-1} \log^2 k$$

und

$$
(8)\qquad \begin{aligned} S_l'' &< a_{22} \sum_{\sqrt{x} < d \leq \sqrt{y}} \frac{1}{d^\alpha}\left(\frac{1}{d^\alpha} + \frac{1}{k}\int_1^{\frac{y}{d}} \frac{du}{u^\alpha}\right) \leq \\ &\leq a_{23}\left(x^{\frac{1}{2}-\alpha} + \frac{y^{1-\alpha}}{k(1-\alpha)} \sum_{d \leq \sqrt{y}} \frac{1}{d}\right) < a_{24} k^{2(1-\alpha^2)-1} \log^2 k. \end{aligned}
$$

Aus (7) und (8) folgt Lemma II.

Es sei

$$
(9a)\qquad \Delta(s, \chi_0) = L(s, \chi_0) - \frac{1}{k(s-1)} \sum_{k^2-k < n \leq k^2} \frac{\chi_0(n)}{n^{s-1}}
$$

und für $\chi \neq \chi_0$

$$
(9b)\qquad \Delta(s, \chi) = L(s, \chi).
$$

Lemma III. *Wenn* $\chi \neq \chi_0$, $\sigma \geq \frac{1}{4}$, $|s| \leq 10$, *dann gilt*

$$
\left| \Delta(s, \chi) - \sum_{n \leq k^2} \frac{\chi(n)}{n^s} \right| < a_{25} k^{\frac{1}{2}-2\sigma} \log k.
$$

Beweis: Für $\sigma > 0$ gilt $L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$. Da nach Pólya $\left| \sum_{n=a}^{b} \chi(n) \right| \leq a_{26} \sqrt{k} \log k$ gilt, folgt durch partielle Summation

$$
\left| \sum_{n=k^2+1}^{\infty} \frac{\chi(n)}{n^s} \right| < a_{27} k^{\frac{1}{2}-2\sigma} \log k, \qquad \text{qu. e. d.}
$$

Lemma IV. *Für* $\sigma \geq \frac{1}{4}$, $|s| \leq 10$ *gilt*

$$
\left| \Delta(s, \chi_0) - \sum_{n \leq k^2} \frac{\chi_0(n)}{n^s} \right| \leq a_{28} k^{1-2\sigma}.
$$

Beweis: Die Funktion $\zeta(s, w)$ $(0 < w \leq 1)$ sei für $\sigma > 1$ durch die Formel

$$
\zeta(s, w) = \sum_{n=0}^{\infty} \frac{1}{(n+w)^s}
$$

definiert. Dann ist bekanntlich[15]) für $\sigma \geq \frac{1}{4}$, wenn m eine beliebige positive ganze Zahl bedeutet,

---

[15]) Siehe z. B. E. Landau, *Vorlesungen über Zahlentheorie* (Leipzig, 1927), Bd. II, p. 9—10.

$$\left| \zeta(s, w) - \frac{1}{(s-1)(m+w)^{s-1}} - \sum_{n=0}^{m} \frac{1}{(n+w)^s} \right| \leqq \frac{|s|}{\sigma} \frac{1}{m^\sigma} < \frac{40}{m^\sigma}.$$

Setzen wir $w = \frac{a}{k}$, multiplizieren mit $k^{-s} \chi_0(a)$, so gelangen wir zu

$$(10) \quad \left| \frac{\chi_0(a)\zeta\left(s, \frac{a}{k}\right)}{k^s} - \frac{\chi_0(km+a)}{k(s-1)(km+a)^{s-1}} - \sum_{n=0}^{m} \frac{\chi_0(kn+a)}{(kn+a)^s} \right| < \frac{a_{29}}{(km+a)^\sigma}.$$

Es sei in (10) $m = k-1$, $a = 1, 2, \ldots k$; wenn wir nach a summieren, folgt

$$\left| \frac{1}{k^s} \sum_{a=1}^{k} \chi_0(a) \zeta\left(s, \frac{a}{k}\right) - \frac{1}{k(s-1)} \sum_{k^2-k+1 \leqq n \leqq k^2} \frac{\chi_0(n)}{n^{s-1}} - \sum_{n=1}^{k^2} \frac{\chi_0(n)}{n^s} \right| < a_{30} k^{1-2\sigma},$$

Da[15])

$$\frac{1}{k^s} \sum_{a=1}^{k} \chi_0(a) \zeta\left(s, \frac{a}{k}\right) = L(s, \chi_0),$$

ist Lemma IV nach der Definition von $\Delta(s, \chi_0)$ (s. (9a)) bewiesen.

Es sei

$$(11) \qquad f(s, \chi) = \Delta(s, \chi) \sum_{n \leqq x} \frac{\chi(n)\mu(n)}{n^s},$$

wo x dieselbe Bedeutung hat, wie in Lemma I. Ferner sei

$$(12) \qquad \sum_{\substack{d \mid n \\ d \leqq x}} \mu(d) = a_n.$$

Dann ist offenbar

$$(13) \qquad |a_n| \leqq d(n),$$

wo $d(n)$ wieder die Teileranzahl von n bedeutet. Wir bemerken schon hier, daß die Zahlen a_n reell und von den Charakteren unabhängig sind.

Lemma V. *Es gilt für* $\sigma \geqq \frac{1}{4}$, $|s| \leqq 10$

$$\left| f(s, \chi_0) - 1 - \sum_{x \leqq n \leqq y} \frac{a_n \chi_0(n)}{n^s} \right| \leqq a_{31} k^{1-2\sigma} \int_1^x \frac{du}{u^\sigma}$$

und für $\chi \neq \chi_0$

$$\left| f(s,\chi) - 1 - \sum_{x \leq n \leq y} \frac{a_n \chi(n)}{n^s} \right| \leq a_{31} k^{\frac{1}{2}-2\sigma} \log k \int_1^x \frac{du}{u^\sigma}.$$

Beweis: Wir werden Lemma V für $\chi \neq \chi_0$ beweisen; für $\chi = \chi_0$ ist der Beweis analog. Aus (11) und Lemma III folgt

$$(14) \quad \left| f(s,\chi) - \sum_{n \leq k^2} \frac{\chi(n)}{n^s} \sum_{n \leq x} \frac{\chi(n)\mu(n)}{n^s} \right| \leq$$

$$\leq a_{32} k^{\frac{1}{2}-2\sigma} \log k \sum_{n \leq x} \frac{1}{n^\sigma} \leq a_{33} k^{\frac{1}{2}-2\sigma} \log k \int_1^x \frac{du}{u^\sigma}.$$

Das linksstehende Produkt im (14) ist gleich

$$\sum_{\substack{n \leq k^2 \\ m \leq x}} \frac{\chi(n)\chi(m)\mu(m)}{(nm)^s} = \sum_{N \leq k^2 x} \frac{\chi(N)}{N^s} \sum_{\substack{m \mid N \\ m \leq x}} \mu(m) = \sum_{N=1}^{k^2 x} \frac{a_N \chi(N)}{N^s}.$$

Da offenbar $k^2 x = y$ und $a_N = 1$ für $N = 1$ und $a_N = 0$ für $1 < N \leq x$ gilt, ist Lemma V bewiesen.

Lemma VI. *Wenn* $\frac{1}{2} \leq \alpha \leq 1 - \frac{2}{\log k}$ *und* $|s| \leq 10$ *ist, dann gilt für* $\sigma \geq \alpha - \frac{1}{\log k}$

$$\left| \sum_{\chi} (f(s,\chi) - 1) \right| < a_{34} k^{2-2\alpha^2} \log^2 k.$$

Beweis: Es bezeichne U_1 die linksseitige Summe. Aus der bekannten Relation $\sum_{\chi} \chi(n) = 0$ oder $= \varphi(k)$ (je nachdem $n \not\equiv$ oder $\equiv 1 \bmod k$) und Lemma V folgt

$$\left| U_1 - \varphi(k) \sum_{\substack{x \leq n \leq y \\ n \equiv 1}} \frac{a_n}{n^s} \right| < a_{31} \left(k^{1-2\alpha} \int_1^x \frac{du}{u^\alpha} + k^{\frac{3}{2}-2\alpha} \int_1^x \frac{du}{u^\alpha} \right) <$$

$$< 2 a_{31} k^{\frac{3}{2}-2\alpha} \cdot k^{2\alpha(1-\alpha)} \log k,$$

also nach (13)

$$(15) \qquad |U_1| \leq k \sum_{\substack{x \leq n \leq y \\ n \equiv 1}} \frac{d(n)}{n^\alpha} + 2 a_{31} k^{\frac{3}{2}-2\alpha^2} \log k.$$

Aus Lemma II und (15) folgt offenbar Lemma VI.

Lemma VII. *Wenn* $\frac{1}{2} \leqq \alpha \leqq 1 - \frac{2}{\log k}$ *und* $|s| \leqq 10$, *so gilt für* $\sigma \geqq \alpha - \frac{1}{\log k}$

$$U_2 = \sum_{\chi} |f(s, \chi) - 1|^2 < a_{32} k^{4(1-\alpha^2)} \log^4 k.$$

Beweis: Aus Lemma V folgt wegen der Ungleichung $|a+b|^2 \leqq 2(|a|^2 + |b|^2)$

$$(16) \qquad U_2 < 2 \sum_{\chi} \left| \sum_{x \leqq n \leqq y} \frac{a_n \chi(n)}{n^s} \right|^2 + a_{33} k^{2-4\alpha} \left(\int_1^x \frac{du}{u^\alpha} \right)^2 \log^2 k.$$

Das erste Glied von (16) ist nach der Orthogonaleigenschaft der Charaktere gleich

$$(17) \qquad \begin{aligned} \sum_{\chi} \sum_{\substack{m,\, n \\ x \leqq m \leqq y \\ x \leqq n \leqq y}} \frac{a_m a_n \chi(n) \bar{\chi}(m)}{(mn)^\sigma} \left(\frac{m}{n}\right)^{it} &= \varphi(k) \sum_{\substack{m \equiv n \\ x \leqq n \leqq y \\ x \leqq m \leqq y \\ (n,\, k) = 1}} \frac{a_m a_n}{(mn)^\sigma} \left(\frac{m}{n}\right)^{it} = \\ = \varphi(k) \sum_{\substack{l \\ (l,\, k)=1}} \left| \sum_{\substack{m \equiv l \\ x \leqq m \leqq y}} \frac{a_m}{m^s} \right|^2 &\leqq \varphi(k) \sum_{l=1}^{k} \left(\sum_{\substack{n \equiv l \\ x \leqq n \leqq y}} \frac{d(n)}{n^\sigma} \right)^2 \end{aligned}$$

nach den bei der Ungleichung (13) gemachten Bemerkungen. Aus (16), (17) und Lemma II folgt

$$U_2 < a_{34} [k^{4(1-\alpha^2)} \log^4 k + k^{2-4\alpha} k^{4\alpha(1-\alpha)} \log^4 k] < a_{35} k^{4(1-\alpha^2)} \log^4 k, \text{ qu.e.d.}$$

Lemma VIII. *Es sei* $\frac{1}{2} \leqq \alpha \leqq 1 - \frac{2}{\log k}$ *und* $F(s)$ *regulär für* $\alpha - \frac{1}{\log k} \leqq \sigma \leqq 4$, $|t| \leqq 10$. *Dann ist die Anzahl der Wurzeln von* $F(s)$ *für* $\alpha \leqq \sigma \leqq \frac{7}{4}$, $|t| \leqq 5$, *wenn nur* $k > a_{36}$, *kleiner als*

$$33 \log^{\frac{3}{2}} k \left[\max_{\substack{\alpha - \frac{1}{\log k} \leqq \sigma \leqq 4 \\ |t| \leqq 7}} \log |F(s)| - \min_{|t| \leqq 5} \log \left| F\left(\frac{3}{2} + it\right) \right| \right].$$

Beweis: Es sei $s_\nu = \frac{3}{2} + i \frac{\nu}{\sqrt{\log k}}$, $(\nu = 0, \pm 1, \pm 2, \ldots, \pm N)$, wo N diejenige ganze Zahl bedeutet, für welche $\frac{N}{\sqrt{\log k}} \leqq 5 < \frac{N+1}{\sqrt{\log k}}$ gilt; dann ist $N \leqq 5 \sqrt{\log k}$. Nach der Jensenschen Formel folgt für die Anzahl der Wurzeln von $F(s)$ im Kreise $|s-s_\nu|\leqq \frac{3}{2}-\alpha+\frac{1}{2\log k}=r$ bei festen ν kleiner als

$$\frac{1}{\log\left(1+\frac{1}{2r\log k}\right)}\left[\frac{1}{2\pi}\int_0^{2\pi}\log\left|F\left(s_\nu+\frac{3}{2}-\alpha+\frac{1}{2\log k}e^{i\vartheta}\right)\right|d\vartheta-\right.$$

$$\left.-\log|F(s_\nu)|\right|<$$

$$<3\log k\left[\max_{|s-s_\nu|\leqq\frac{3}{2}-\alpha+\frac{1}{\log k}}\log|F(s)|-\min_{|t|\leqq 5}\log\left|F\left(\frac{3}{2}+it\right)\right|\right]$$

Wenn wir nach ν summieren, folgt, daß die Anzahl derjenigen Wurzeln von $F(s)$, welche in irgendwelche der obigen $2N+1$ Kreise fallen, kleiner als

$$33\log^{\frac{3}{2}}k\left[\max_{\substack{\alpha-\frac{1}{\log k}\leqq\sigma\leqq 4\\ |t|\leqq 7}}\log|F(s)|-\min_{|t|\leqq 5}\log\left|F\left(\frac{3}{2}+it\right)\right|\right]$$

ist. Da diese Kreise offenbar das Gebiet $\alpha\leqq\sigma\leqq\frac{7}{4}$, $|t|\leqq 5$ bedecken, folgt die Behauptung.

Beweis des Satzes I. Für $\alpha\leqq\sigma\leqq\frac{7}{4}$, $|t|\leqq 5$ ist $L(s,\chi_0)\neq 0$; nach der Definition der Funktionen $f(s,\chi)$ ist also $N_k(\alpha)$ gewiß nicht größer, als die Gesamtanzahl der hier liegenden Wurzeln aller Funktionen $f(s,\chi)$. Lemma VIII ergibt mit $F(s)=\prod_\chi f(s,\chi)$ die Ungleichung

$$(18)\quad N_k(\alpha)<33\log^{\frac{3}{2}}k\left[\frac{1}{2}\max_{\substack{\alpha-\frac{1}{\log k}\leqq\sigma\leqq 4\\ |t|\leqq 7}}\sum_\chi\log|f(s,\chi)|^2-\right.$$

$$\left.-\min_{\substack{|t|\leqq 5\\ \sigma=\frac{3}{2}}}\sum_\chi\log|f(s,\chi)|\right].$$

Da aber für $\sigma=\frac{3}{2}$ und $\chi\neq\chi_0$

$$f(s,\chi)=\left(\sum_{n=1}^{\infty}\frac{\chi(n)}{n^s}\right)\left(\sum_{m\leqq x}\frac{\chi(m)\mu(m)}{m^s}\right)=$$

$$=\sum_{N=1}^{\infty}\frac{\chi(N)}{N^s}\left(\sum_{\substack{d\leqq x\\ d|N}}\mu(d)\right)=1+\sum_{N=[x]+1}^{\infty}\frac{\chi(N)}{N^s}a_N,$$

also

$$|f(s,\chi)-1| < \sum_{N=[x]+1}^{\infty} \frac{d(N)}{N^{3/2}} < a_{37}\frac{\log k}{k^{\alpha}}$$

und da analoges gilt für $\chi=\chi_0$, also

$$\sum_{\chi} \log\left|f\left(\frac{3}{2}+it,\chi\right)\right| > -a_{37}k^{1-\alpha}\log k,$$

so folgt aus (18) nach dem Satz über das arithmetische und geometrische Mittel

$$N_k(\alpha) < 17\log^{\frac{3}{2}} k\left[\varphi(k) \max_{\substack{\alpha-\frac{1}{\log k}\leqq\sigma\leqq 4\\ |t|\leqq 7}} \log\left(\frac{1}{\varphi(k)}\sum_{\chi}|f(s,\chi)|^2\right)+\right.$$
$$\left.+a_{38}k^{1-\alpha}\log k\right] <$$
$$< a_{39}\log^{\frac{3}{2}} k\left[\varphi(k)\max_{\substack{\alpha-\frac{1}{\log k}\leqq\sigma\leqq 4\\ |t|\leqq 7}}\log\left(1+\frac{2}{\varphi(k)}\Re\sum_{\chi}(f(s,\chi)-1)+\right.\right.$$
$$\left.\left.+\frac{1}{\varphi(k)}\sum_{\chi}|f(s,\chi)-1|^2\right)+k^{1-\alpha}\log k\right] <$$
$$< a_{40}\left[k^{1-\alpha}\log^{\frac{5}{2}} k+\log^{\frac{3}{2}} k \max_{\substack{\alpha-\frac{1}{\log k}\leqq\sigma\leqq 4\\ |t|\leqq 7}}\left|\sum_{\chi}(f(s,\chi)-1)\right|+\right.$$
$$\left.+\log^{\frac{3}{2}} k \max_{\substack{\alpha-\frac{1}{\log k}\leqq\sigma\leqq 4\\ |t|\leqq 7}} \sum_{\chi}|f(s,\chi)-1|^2\right].$$

Dann ist aber nach Lemma VI und VII

$$N_k(\alpha) < a_{41}k^{4(1+\alpha)(1-\alpha)}\log^{\frac{11}{2}} k, \qquad \text{qu. e. d.}$$

Der Beweis des Satzes II verläuft analog, wir skizzieren ihn nur daher. Es sei $1-\frac{2}{\log k} \geqq \alpha \geqq \frac{1}{2}+\frac{5}{\sqrt{\log\log k}}$ und

(19) $$\Delta_1(s,\chi_0) = L(s,\chi_0) - \frac{1}{k(s-1)}\sum_{n=1}^{k}\frac{\chi_0(n)}{n^{s-1}}$$

und für $\chi \neq \chi_0$ wieder

(20) $$\Delta_1(s,\chi) = L(s,\chi).$$

Die Rolle des Lemmas III bzw. IV übernehmen die Ungleichungen

$$\left|\Delta_1(s,\chi)-\sum_{n\leqq k}\frac{\chi(n)}{n^s}\right|<a_{42}k^{\frac{1}{2}-\sigma}\log k \tag{21}$$

für $\chi\neq\chi_0$, $\sigma\geqq\frac{1}{4}$, $|s|\leqq 10$ und

$$\left|\Delta_1(s,\chi_0)-\sum_{n\leqq k}\frac{\chi_0(n)}{n^s}\right|<a_{43}k^{1-\sigma}. \tag{22}$$

Es sei

$$f_1(s,\chi)=\Delta_1(s,\chi)\sum_{n\leqq z}\frac{\chi(n)\,\mu(n)}{n^s},$$

wo z zur Abkürzung für $k^{2\alpha-1-\frac{2}{\log k}}$ steht. Statt Lemma V hat man für $\sigma\geqq\frac{1}{4}$, $|s|\leqq 10$

$$\left|f_1(s,\chi_0)-1-\sum_{z\leqq n\leqq kz}\frac{a_n\chi_0(n)}{n^s}\right|<a_{44}k^{1-\sigma}\int_1^z\frac{du}{u^\sigma} \tag{23}$$

und für $\chi\neq\chi_0$

$$\left|f_1(s,\chi)-1-\sum_{z\leqq n\leqq kz}\frac{a_n\chi(n)}{n^s}\right|\leqq a_{45}k^{\frac{1}{2}-\sigma}\log k\int_1^z\frac{du}{u^\sigma}. \tag{24}$$

Aus (23) und (24) folgt durch Summation

$$\left|\sum_\chi(f_1(s,\chi)-1)-\varphi(k)\sum_{\substack{z\leqq n\leqq kz\\ n\equiv 1}}\frac{a_n}{n^s}\right|<a_{46}k^{\frac{3}{2}-\sigma}\log k\int_1^{k^{2\alpha-1}}\frac{du}{u^\sigma},$$

$$\left|\sum_\chi(f_1(s,\chi)-1)\right|\leqq\varphi(k)\sum_{z\leqq n\leqq kz,\ n\equiv 1}\frac{d(n)}{n^\sigma}+a_{47}k^{\frac{3}{2}-\sigma}\log k\int_1^{k^{2\alpha-1}}\frac{du}{u^\sigma}. \tag{25}$$

Da bekanntlich $d(n)\leqq a_{48}\exp\left(\dfrac{\frac{5}{4}\log n}{\log\log n}\right)$, so folgt aus (25)

$$\left|\sum_\chi(f_1(s,\chi)-1)\right|\leqq a_{49}\left[k^{(1-2\alpha)\sigma+1}+\int_1^{k^{2\alpha}}\frac{du}{u^\sigma}\right]\exp\left(3\frac{\log k}{\log\log k}\right)+ \tag{26}$$

$$+a_{50}k^{\frac{3}{2}-\sigma}\log k\int_1^{k^{2\alpha-1}}\frac{du}{u^\sigma}.$$

Wenn also

$$(27)\qquad \frac{1}{2}+\frac{3}{\log\log k}\leqq \alpha \leqq \frac{-1+\sqrt{8}}{2}\ (<1),$$

so folgt aus (26) für $\sigma \geqq \alpha - \dfrac{1}{\log k}$

$$\left|\sum_{\chi}(f_1(s,\chi)-1)\right| < a_{51}\left\{\left[k^{1-(2\alpha-1)\alpha}+k^{2\alpha(1-\alpha)}\right]\exp\left(\frac{3\log k}{\log\log k}\right)+\right.$$
$$\left.+k^{\frac{3}{2}-\alpha}\log k\,.\,k^{(2\alpha-1)(1-\alpha)}\right\}.$$

Da $1-(2\alpha-1)\alpha > 2\alpha(1-\alpha)$ und $\dfrac{3}{2}-\alpha+(2\alpha-1)(1-\alpha) >$ $> 1-(2\alpha-1)\alpha+\dfrac{3}{\log\log k}$, so gilt

$$(28)\qquad \left|\sum_{\chi}(f_1(s,\chi)-1)\right| < a_{52}\,k^{\frac{1}{2}+2\alpha-2\alpha^2}\log k.$$

Ferner gilt, wie bei Lemma VII und Lemma I, für $\sigma \geqq \alpha - \dfrac{1}{\log k}$,

$$\sum_{\chi}|f_1(s,\chi)-1|^2 < 2\varphi(k)\sum_{l=1}^{k}\Bigg(\sum_{\substack{z\leqq n\leqq kz\\ n\equiv l}}\frac{d(n)}{n^{\alpha}}\Bigg)^2+$$
$$+c_{53}\,k^{2-2\alpha}\log^2 k\left[\int_1^z\frac{du}{u^{\alpha}}\right]^2<$$
$$(29)\qquad < c_{54}\left\{k\exp\left(\frac{6\log k}{\log\log k}\right)\sum_{l=1}^{k}\Bigg(\sum_{\substack{z\leqq n\leqq kz\\ n\equiv l}}\frac{1}{n^{\alpha}}\Bigg)^2+\right.$$
$$\left.+k^{2-2\alpha}\log^2 k\,.\,k^{2(2\alpha-1)(1-\alpha)}\right\} < c_{56}\,k^{4\alpha(1-\alpha)}\exp\left(\frac{6\log k}{\log\log k}\right).$$

Aus (28), (29) und Lemma VIII folgt Satz II ebenso, wie Satz I, da für $\dfrac{1}{2}+\dfrac{5}{\sqrt{\log\log k}}\leqq\alpha\leqq\dfrac{-1+\sqrt{8}}{2}$ und $k > c_{57}$

$$\frac{1}{2}+2\alpha-2\alpha^2 > 4\alpha(1-\alpha)+\frac{6}{\log\log k},$$

also

$$\left|\sum_{\chi}(f_1(s,\chi)-1)\right|+\sum_{\chi}|f_1(s,\chi)-1|^2 < c_{58}\,k^{\frac{1}{2}+2\alpha-2\alpha^2}\log k.$$

(Eingegangen am 27. August 1941.)

Über Kontraktionen des Hilbertschen Raumes.

Von FRIEDRICH RIESZ und BÉLA V. SZ. NAGY in Szeged.

Der folgende, sog. statistische Ergodensatz wurde von J. v. NEUMANN in 1931 gefunden[1]):

Ist U eine unitäre Transformation des Hilbertschen Raumes $\mathfrak{H}$,[2]) *so existiert für jedes* $f \in \mathfrak{H}$ *der Limes*

$$f^* = \lim_{n-m\to\infty} \frac{1}{n-m} \sum_{\nu=m}^{n-1} U^\nu f,$$

und f^* ist invariant bei U, d. h. $Uf^* = f^*$.

Der folgende Beweis dieses Satzes (und einer Verallgemeinerung desselben), den einer von uns (F. RIESZ) vor fast 10 Jahren gefunden hat, führte zu einer Frage, die am Ende von §. 1 formuliert und in §. 2 (durch B. v. SZ. NAGY) beantwortet wird.

Genauer gesagt, gab CARLEMAN in 1932 einen sehr elementaren Beweis dieses Satzes, allerdings nur für den sogenannten ergodischen Fall, d. h. für den Fall, wo der Hilbertsche Raum von den quadratisch integrierbaren Funktionen gebildet wird und

[1]) J. v. NEUMANN, Proof of the quasi-ergodic hypothesis, *Proceedings National Academy U. S. A.*, 18 (1932), S. 70—82.

[2]) Für die Grundbegriffe der Theorie des Hilbertschen Raumes verweisen wir auf M. H. STONE, *Linear transformations in Hilbert space* (New York, 1932), oder auf B. v. Sz. NAGY, Spektraldarstellung linearer Transformationen des Hilbertschen Raumes, *Ergebnisse der Math. und ihrer Grenzgebiete*, Bd. V/5 (Berlin, 1942). Der Verfasser des letztgenannten Berichtes benutzt die Gelegenheit, einen infolge einer Textänderung während der Korrektur eingeschlichenen Defekt richtigzustellen:

Auf Seite 17, Reihen 7—9, soll der Satz: „Umgekehrt, es folgt aus $P \geq Q$, daß ... $PQ = Q$." *richtig heißen:* „Umgekehrt, es folgt aus $P \geq Q$, daß $I - P \leq I - Q$; für jedes f gilt also $\|(I-P)Qf\|^2 = ((I-P)Qf, Qf) \leq$ $\leq ((I-Q)Qf, Qf) = 0$, d. h. $(I-P)Q = O$, $Q = PQ$."

die einzigen Invarianten von U die Konstanten sind[3]). F. RIESZ hat nun bemerkt, daß der Gedanke von CARLEMAN, passend umgestaltet, nicht nur den ursprünglichen Satz, sondern unter anderem auch dessen Verallgemeinerung auf den Fall einer beliebigen normalen Transformation T liefert, wenn nur T eine *Kontraktion* ist, d. h. wenn $\|Tf\| \leqq \|f\|$ für jedes $f \in \mathfrak{H}$ gilt. Ja auch die Voraussetzung, daß T normal ist, dient nur dazu, zu sichern, daß jedes Element, das bei T^* invariant ist, auch bei T invariant bleibt.

Den so verallgemeinerten Gedankengang hat RIESZ im Druck nicht veröffentlicht, dieser wurde nur mündlich und brieflich weitergegeben, und gelangte schließlich, wieder nur für den Fall eines unitären U, im Berichte von E. HOPF über Ergodentheorie[4]) zum Abdruck. In dieser Form lautet der Beweis, wie folgt.

§. 1.

Sei U eine unitäre Transformation des Hilbertschen Raumes $\mathfrak{H}$. Der von allen Elementen der Form $Ug-g$ aufgespannte Unterraum von $\mathfrak{H}$ sei mit $\mathfrak{M}$ bezeichnet. Ein Element f aus $\mathfrak{M}$ ist dann entweder selbst von dieser Form, oder es ist mindestens durch solche Elemente beliebig genau approximierbar.

Wir setzen

$$V_{mn} = \frac{1}{n-m} \sum_{\nu=m}^{n-1} U^\nu \qquad (n > m \geqq 0).$$

Ist $f = Ug-g$, so hat man

$$V_{mn} f = \frac{U^n g - U^m g}{n-m},$$

$$(1) \qquad \|V_{mn} f\| \leqq \frac{\|U^n g\| + \|U^m g\|}{n-m} = \frac{2\|g\|}{n-m},$$

folglich $V_{mn} f \to 0$, wenn $n-m \to \infty$.

Ist f zwar nicht selbst von dieser Form, aber mit solchen Elementen $f' = Ug-g$ beliebig genau approximierbar, so folgt aus

---

[3]) T. CARLEMAN, Application de la théorie des équations intégrales linéaires aux équations différentielles non linéaires, *Acta Math.*, **59** (1932), S. 63-87.

[4]) E. HOPF, Ergodentheorie, *Ergebnisse der Math. und ihrer Grenzgebiete*, Bd. V/2 (Berlin, 1937), § 8.

$$(2)\quad \|V_{mn}f\| \leqq \|V_{mn}(f-f')\| + \|V_{mn}f'\| \leqq$$
$$\leqq \frac{1}{n-m}\sum_{\nu=m}^{n-1}\|U^{\nu}(f-f')\| + \|V_{mn}f'\| = \|f-f'\| + \|V_{mn}f'\|$$

wieder $V_{mn}f \to 0$ für $n-m \to \infty$.

Der Satz ist also für jedes f aus $\mathfrak{M}$ gültig. Andererseits gilt aber der Satz für jedes Element, das bei U invariant ist, da dies auch bei V_{mn} invariant ist. Diese Invarianten bilden einen Unterraum $\mathfrak{N}$ von $\mathfrak{H}$.

Wegen der Linearität von V_{mn} gilt dann der Satz auch für jedes Element $f = f_1 + f_2$ mit $f_1 \in \mathfrak{M}$, $f_2 \in \mathfrak{N}$, und f^* ist dann gleich f_2, d. h. ist bei U invariant.

Nun ist jedes Element f von $\mathfrak{H}$ in dieser Form darstellbar. Es genügt dazu zu zeigen, daß jedes Element g, das zu $\mathfrak{M}$ orthogonal steht, in $\mathfrak{N}$ enthalten ist. In der Tat, für beliebiges $h \in \mathfrak{H}$ gilt:

$$(U^*g - g, h) = (U^*g, h) - (g, h) = (g, Uh) - (g, h) = (g, Uh - h) = 0,$$

folglich ist $U^*g - g = 0$, d. h. g ist gegenüber U^* invariant. Auf Grund der für die unitären Transformationen geltenden Gleichung $UU^* = I$ folgt hieraus $Ug = U(U^*g) = g$, d. h. g ist auch gegenüber U invariant, also in $\mathfrak{N}$ enthalten.

Damit ist der Satz bewiesen. Von der Voraussetzung, daß U unitär ist, wurde nur benutzt, daß $\|Uf\| = \|f\|$ (letzter Schritt in (1) und (2)), und daß jedes Element, das bei U^* invariant ist, auch bei U invariant bleibt. *Der Gedankengang gilt also umsomehr für jede solche beschränkte lineare Transformation T statt U,*

a) *die eine Kontraktion ist, d. h. für die immer $\|Tf\| \leqq \|f\|$ gilt,*

b) *die mit ihrer Adjungierten T^* die gleichen Elemente invariant läßt.*

In den letzten Jahren gelang es, den v. Neumannschen Satz durch verschiedene sehr einfache Methoden derart zu verallgemeinern, daß die Rolle von U eine beliebige Kontraktion, d. h. eine nur die Bedingung a) genügende lineare Transformation T übernimmt[5]). Das legt die Frage nahe, ob Bedingung b) nicht

[5]) Vgl. z. B. Garrett Birkhoff, The mean ergodic theorem, *Duke Math. Journal*, **5** (1939), S. 19—20; E. R. Lorch, Means of iterated transformations in reflexive vector spaces, *Bulletin American Math. Society*, **45** (1939), S. 945—947; F. Riesz, Some mean ergodic theorems, *Journal London Math. Society*, **13** (1938), S. 274—278; Another proof of the mean ergodic theorem, *diese Acta*, **10** (1941), S. 75—76.

schon aus Bedingung a) folgt? Wir werden diese Frage bejahend beantworten.

§. 2.

Wir beweisen den soeben angekündigten Satz:

Ist T eine Kontraktion, d. h. eine lineare Transformation mit $\|Tf\| \leqq \|f\|$ für jedes f aus $\mathfrak{H}$, so haben T und ihre Adjungierte T^* die gleichen invarianten Elemente.

Mit T ist bekanntlich auch T^* eine Kontraktion, d. h. es gilt auch

$$\|T^* f\| \leqq \|f\|.$$

Schreibt man hier insbesondere $f = Tg$, so erhält man

$$\|T^* Tg\| \leqq \|Tg\|. \tag{3}$$

Man hat ferner

$$\begin{aligned}\|g - T^* Tg\|^2 &= (g - T^* Tg, g - T^* Tg) = \\ &= (g, g) - (g, T^* Tg) - (T^* Tg, g) + (T^* Tg, T^* Tg) = \\ &= \|g\|^2 - 2\|Tg\|^2 + \|T^* Tg\|^2,\end{aligned}$$

also, wegen (3),

$$\|g - T^* Tg\|^2 \leqq \|g\|^2 - \|Tg\|^2. \tag{4}$$

Nun sei g ein gegenüber T invariantes Element. Dann ist die rechte Seite von (4) gleich 0, folglich ist $g - T^* Tg = 0$, d. h., g ist auch gegenüber $T^* T$ invariant.

Also ist

$$T^* g = T^*(Tg) = g,$$

d. h., g ist auch gegenüber T^* invariant.

Damit wurde gezeigt, daß die Elemente, die bei T invariant sind, auch bei T^* invariant bleiben. Da man aber die Rolle von T und T^* miteinander vertauschen kann, folgt hieraus, daß T und T^* dieselben Invarianten besitzen.

(Eingegangen am 4. Mai 1942.)

Über orthogonale Entwicklungen.

Von S. SIDON (†) in Budapest.

Die hier zu behandelnden Fragen sind sehr verschiedenartig und beziehen sich teils auf allgemeine orthogonale, teils auf Fourierschen und Walshsche Entwicklungen. Die Einteilung der Sätze geschieht nach dem Charakter derselben in vier Gruppen: im §. 1 werden Absolut-Konvergenzfragen, im §. 2 einige Interpolationsfragen, im §. 3 Fragen aus dem Ideenkreis des Young—Hausdorffschen Satzes, im §. 4 eine Frage von STEINHAUS und anknüpfende Sätze behandelt. Die Sätze der §§. 1 und 3 haben natürlich einige Berührungspunkte. Die Besprechung der Resultaten führen wir in den einzelnen Paragraphen aus. Einige Sätze, welche vereinzelnt stehen, folgen im Anhang.

§. 1.

Eine Reihe

(1) $$\sum_{\nu=0}^{\infty}(a_\nu \cos \nu x + b_\nu \sin \nu x)$$

heißt bekanntlich „stark lakunär", wenn diejenige Indizes $\nu = n_k$, für welche $a_\nu^2 + b_\nu^2 > 0$, „sehr selten" liegen, d. h. es gibt ein $q > 1$ so, daß

(2) $$\frac{n_{k+1}}{n_k} \geqq q, \qquad k = 1, 2, \ldots$$

(†) Der durch seine Arbeiten über Orthogonalreihen wohlbekannte Mathematiker S. SIDON ist am 27. April 1941 nach einem leichten Unfall an Lungenentzündung gestorben. Seine hier zu Veröffentlichung kommende Arbeit wurde aus kurzen Manuskripten zusammengestellt, die im Zeitintervall 5. Juni 1937—13. Juni 1940 bei unserer Redaktion eingegangen sind. Die meisten derselben bedürften aus Lesbarkeitsrücksichten noch einer Umarbeitung. Die vorliegende Umarbeitung verdankt die Redaktion der Herren G. GRÜNWALD und P. TURÁN.

Eine solche Reihe können wir im Gestalt

$$(3)\qquad \sum_{\nu=0}^{\infty}(a_\nu \cos n_\nu x + b_\nu \sin n_\nu x)$$

schreiben, da wir Summabilitätsfragen nicht berühren. Wenn $f(x)$ eine einseitig beschränkte und L-integrable Funktion bedeutet, deren Fourier-Reihe stark lakunär ist, so habe ich bewiesen[1]), daß die Koeffizientenreihe absolut konvergiert. Mit einer Umarbeitung der Beweisideen beweise ich hier zunächst den

Satz 1. *Es sei $f(x)$ wieder einseitig beschränkt und L-integrabel und wir setzen voraus, daß ihre Fourierreihe in k stark lakunäre trigonometrische Reihen zerfällt, d. h.*

$$(4)\qquad f(x) \sim \sum_{i=1}^{k}\sum_{j=1}^{\infty}(a_{ji}\cos n_{ji}x + b_{ji}\sin n_{ji}x),$$

wo die Zahlen n_{ji} alle verschieden sind und mit einem $q > 1$

$$(5)\qquad \frac{n_{j+1,1}}{n_{j1}} > q,\ \frac{n_{j+1,2}}{n_{j2}} > q, \ldots, \frac{n_{j+1,k}}{n_{jk}} > q$$

gilt. Dann ist $\sum_{i=1}^{k}\sum_{j=1}^{\infty}(|a_{ji}| + |b_{ji}|) < \infty.$

Beweis. Ohne Beschränkung der Allgemeinheit können wir voraussetzen

$$(6)\qquad f(x) \leqq M,\quad \int_0^{2\pi} f(x)\,dx = 0.$$

Es sei $x = x_0$ eine feste, aber beliebige Stelle und

$$(7)\qquad \operatorname{sign}(a_{ji}\cos n_{ji}x_0 + b_{ji}\sin n_{ji}x_0) = \varepsilon_{ji}.$$

Die positive ganze Zahl l sei die kleinste, für welche

$$(8)\qquad 1 + \frac{1}{q^l - 1} < \sqrt{q},\quad 1 - \frac{1}{q^l - 1} > \frac{1}{\sqrt{q}}$$

gilt und wir betrachten die Funktionen

$$(9)\qquad f_{mi}(x) = \sum_{r=0}^{l-1}\prod_{j=0}^{m}\left(1 + \frac{1}{k}\varepsilon_{lj+r,i}\cos n_{lj+r,i}x\right).$$

---

[1]) S. Sidon, Ein Satz über die absolute Konvergenz von Fourierreihen, in denen sehr viele Glieder fehlen, *Math. Annalen*, **96** (1927), S. 418—419; Verallgemeinerung eines Satzes über die absolute Konvergenz von Fourierreihen mit Lücken, *Math. Annalen*, **97** (1927), S. 675—676. Für eine wichtige Anwendung siehe S. Banach, Über einige Eigenschaften der lakunären trigonometrischen Reihen, *Studia Math.*, **2** (1930), S. 207—228.

Wir behaupten zunächst, daß

$$(10)\qquad \frac{1}{2\pi}\int_0^{2\pi} f_{mi}(x)\,dx = l.$$

Das wird bewiesen sein, indem wir zeigen, daß keine der Zahlen von der Form

$$N = n_{i_1 l+r,\,i} \pm n_{i_2 l+r,\,i} \pm \ldots, \qquad i_1 > i_2 > \ldots$$
$$(l,\ r,\ i \text{ fest})$$

verschwindet. Nach (5) und (8) gilt nämlich

$$(11)\qquad \begin{aligned} N &\geqq n_{i_1 l+r,\,i} - n_{i_2 l+r,\,i} - \ldots > n_{i_1 l+r,\,i}\left(1-\frac{1}{q^l}-\frac{1}{q^{2l}}-\ldots\right) > \\ &> n_{i_1 l+r,\,i}\left(1-\frac{1}{q^l-1}\right) > n_{i_1 l+r,\,i}\,\frac{1}{\sqrt{q}} > 0. \end{aligned}$$

Ganz ähnlich sieht man ein, daß man, wenn wir ein beliebiges Produkt aus (9) ausmultiplizieren und jedes Glied nach der Identität

$$\cos x_1 \cos x_2 \ldots \cos x_n = \frac{1}{2^{n-1}} \sum \cos(x_1 \pm x_2 \pm \ldots \pm x_n)$$

umformen, nie zwei Glieder sich zusammenziehen lassen. Wir behaupten aber, daß man auch dann nie zwei Glieder zusammenziehen kann, wenn wir in (9) nach r summieren. Im entgegengesetzten Falle gäbe es nämlich zwei ganzzahlige Wertsysteme $i_1 > i_2 > \ldots > i_s$ und $j_1 > j_2 > \ldots > j_t$ und zwei ganze Zahlen $0 \leqq r_1,\ r_2 \leqq l-1$ so, daß

$$(12)\qquad \begin{aligned} n_{i_1 l+r_1,\,i} \pm n_{i_2 l+r_1,\,i} \pm \ldots \pm n_{i_s l+r_1,\,i} &= \\ &= n_{j_1 l+r_2,\,i} \pm n_{j_2 l+r_2,\,i} \pm \ldots \pm n_{j_t l+r_2,\,i}. \end{aligned}$$

Ohne Beschränkung der Allgemeinheit sei $i_1 l + r_1 \geqq j_1 l + r_2$. Es sei erstens das Zeichen $>$ gültig. Dann ist die linke Seite nach (8) größer als

$$(13\text{a})\qquad n_{i_1 l+r_1,\,i}\left(1-\frac{1}{q^l}-\frac{1}{q^{2l}}-\ldots\right) > n_{i_1 l+r_1,\,i}\,\frac{1}{\sqrt{q}},$$

die rechte Seite aber nach (8) kleiner als

$$(13\text{b})\qquad n_{j_1 l+r_2,\,i}\left(1+\frac{1}{q^l}+\frac{1}{q^{2l}}+\ldots\right) < n_{j_1 l+r_2,\,i}\sqrt{q}$$

und nach (5)

$$(13\text{c})\qquad n_{i_1 l+r_1,\,i}\,\frac{1}{\sqrt{q}} > n_{j_1 l+r_2,\,i}\sqrt{q}.$$

(13a), (13b) und (13c) sind also mit (12) nicht verträglich. Wenn zweitens $i_1 l + r_1 = j_1 l + r_2$, so verläuft der Beweis ähnlich. Also gilt für jedes $1 \leqq i \leqq k$

$$\text{(14)} \qquad \text{Koeff. } \cos n_{\nu i} x \text{ in } f_{mi}(x) = \frac{1}{k} \varepsilon_{\nu i}.$$

Es sei

$$\text{(15)} \qquad F_m(x) = \sum_{i=1}^{k} f_{mi}(x)$$

Nach (10) gilt offenbar

$$\text{(16a)} \qquad \frac{1}{2\pi} \int_0^{2\pi} F_m(x)\, dx = kl$$

und

$$\text{(16b)} \qquad F_m(x) \geqq 0 \qquad 0 \leqq x \leqq 2\pi.$$

Wir betrachten zuerst den Koeffizienten von $\cos n_{j1} x$ in $F_m(x)$. Der Beitrag von $f_{m1}(x)$ ist nach (14) $\frac{1}{k} \varepsilon_{j1}$. Es ist aber möglich, daß z. B. auch in $f_{m2}(x)$ das Glied $\cos n_{j1} x$ auftritt. Nach den obigen hat die Gleichung

$$\text{(17a)} \qquad n_{j_1 l+r,2} \pm n_{j_2 l+r,2} \pm \ldots \pm n_{j_s l+r,2} = n_{j,1} \qquad (j_1 > j_2 > \ldots > j_s)$$

bei veränderlichen $j_1, j_2, \ldots, j_s$ und r höchstens eine Lösung; diese gibt also zu dem Koeffizienten von $\cos n_{j1} x$ den Beitrag

$$\text{(17b)} \qquad \frac{1}{2^{s-1}} \frac{1}{k^s} \varepsilon_{j_1 l+r,2}, \varepsilon_{j_2 l+r,2} \ldots \varepsilon_{j_s l+r,2}.$$

Da offenbar $s \geqq 2$ gilt, ist der Beitrag höchstens $\frac{1}{2k^2}$. Dasselbe gilt offenbar auch für die Beiträge von $f_{m3}(x), \ldots, f_{mk}(x)$. Also „vorwiegt" der Beitrag von $f_{m1}(x)$ beim Koeffizienten von $\cos n_{j1} x$ und zwar gibt

$$\frac{1}{\pi} \int_0^{2\pi} F_m(x) \cos n_{j1} x\, dx - \frac{1}{\pi k} \varepsilon_{j1} = \vartheta_{j1} \frac{1}{2\pi k}, \qquad |\vartheta_{j1}| \leqq 1.$$

Analog gilt

$$\text{(18)} \quad \frac{1}{\pi} \int_0^{2\pi} F_m(x) \cos n_{ji} x\, dx - \frac{1}{\pi k} \varepsilon_{ji} = \vartheta_{ji} \frac{1}{2\pi k}, \qquad |\vartheta_{ji}| \leqq 1$$

$$i = 1, 2, \ldots, k;\ j = 1, 2, \ldots, lm + l - 1.$$

Die Fortsetzung des Beweises verläuft analog, wie ich l. c. [1]) angeführt habe; der Vollständigkeit halber skizzieren wir ihn. Es gilt

$$S = \sum_{i=1}^{k} \sum_{j=0}^{lm+l-1} |a_{ji} \cos n_{ji} x_0 + b_{ji} \sin n_{ji} x_0| \leq$$

$$\leq 2k\pi \sum_{i=1}^{k} \sum_{j=0}^{lm+l-1} \frac{1+\frac{\vartheta_{ji}}{2}}{k\pi} |a_{ji} \cos n_{ji} x_0 + b_{ji} \sin n_{ji} x_0| =$$

$$= 2k\pi \sum_{i=1}^{k} \sum_{j=}^{lm+l-1} \frac{1+\frac{\vartheta_{ji}}{2}}{k\pi} \varepsilon_{ji}(a_{ji} \cos n_{ji} x_0 + b_{ji} \sin n_{ji} x_0),$$

also nach (18), (6), (16b) und (16a)

$$(19) \qquad \begin{aligned} S &\leq 2k \int_0^{2\pi} f(x) F_m(x-x_0)\,dx < 2kM \int_0^{2\pi} |F_m(x-x_0)|\,dx = \\ &= 2kM \int_0^{2\pi} F_m(x)\,dx = 2k^2 l\pi M \end{aligned}$$

unabhängig von x_0 und m. Daraus folgt, daß die Reihe

$$\sum_{j,i} |a_{ji} \cos n_{ji} x_0 + b_{ji} \sin n_{ji} x_0|$$

überall konvergiert, also gilt dasselbe auch für die Reihe $\sum_{ji} (|a_{ji}| + |b_{ji}|)$ selbst, w. z. b. w.

Der soeben bewiesene Satz läßt sich unmittelbar auf fastperiodische Reihen $\sum_{k=1}^{\infty} (a_k \cos \lambda_k x + b_k \sin \lambda_k x)$ mit $\lambda_k > 0$ ausdehnen. Die Bedingung (5) kann dann bei sämtlichen Exponentenfolgen λ_{ki} oder einen Teil derselben durch die der linearen Unabhängigkeit der Glieder der nämlichen Exponentenfolgen ersetzt werden.

Es ist naheliegend zu fragen, ob der Satz sich nicht dahin verschärfen läßt, daß statt der Voraussetzung $\frac{n_{k+1}}{n_k} > q > 1$ auch $\sqrt[k]{n_k} > 1$ genügt. Dann gilt aber kein Satz über absoluter Konvergenz; auch dann nicht, wenn wir statt der einseitigen Beschränktheit überall Stetigkeit voraussetzen. Diese Tatsache folgt fast unmittelbar aus den bekannten Fejérschen Konstruktionsverfahren einer überall stetigen Funktion mit divergenter Fourierreihe[2]).

---

[2]) L. Fejér, Sur les singularités des séries de Fourier des functions continues, *Annales de l'École Normale Supérieure*, 28 (1911), S. 63–103.

Im Folgenden werden wir auch einige Sätze über die Rademacherschen und Walshschen Orthogonalsysteme entwickeln. Das Rademachersche System besteht bekanntlich aus den Funktionen

(20a) $\varphi_n(x) = \operatorname{sign} \sin 2^{n+1} x \qquad 0 \leqq x \leqq 1, \; n = 0, 1, 2, \ldots$

und das Walshsche System aus den Funktionen

(20b) $\psi_n(x) = \varphi_{n_1}(x) \varphi_{n_2}(x) \ldots \varphi_{n_\nu}(x) \qquad \psi_0(x) = 1, \; n = 1, 2, \ldots,$

wo $\varphi_{n_1}(x), \ldots$ die Bedeutung (20a) haben und die Entwicklung von n im dyadischen System

(20c) $n = 2^{n_1} + 2^{n_2} + \ldots + 2^{n_\nu}, \qquad n_1 > n_2 > \ldots > n_\nu \geqq 0$

lautet. Wenn n die Darstellung (20c) besitzt, werden wir sagen, daß die Zahl n ν-ziffrig ist.

Das Rademachersche System wurde schon hinsichtlich lakunärer Eigenschaften untersucht[3]); nicht aber des Walshsche System. Für diesen gilt

S a t z II. *Wenn die Walsh-Entwicklung einer einseitig beschränkten und L-integrablen $f(x)$ stark lakunär ist, so ist die Reihe der Koeffizienten absolut konvergent.*

Der Beweis verläuft ganz analog, wie der des analogen Satzes beim trigonometrischen System[1]).

Bei einer anderen Verallgemeinerung des Satzes setzen wir von der gewöhnlichen Fourierreihe der einseitig beschränkten und L-integrablen $f(x)$ nicht mehr starke Lakunarität voraus, sondern nur, daß unendlich viele Indizes $n = n_k$ existieren, in deren „Nähe" alle übrigen Koeffizienten verschwinden. Dann ist die Reihe der „einsamen" Koeffizienten absolut konvergent. Die Voraussetzung dieses Satzes ist wohl nicht unnatürlich; die Untersuchungen von OSTROWSKI[4]) zeigen, wie wichtige Rolle eine ähnliche Lakunaritätsbedingung beim Überkonvergenz spielt. Genauer gesagt, beweisen wir den folgenden

S a t z III. *Es sei $f(x)$ einseitig beschränkt L-integrierbar und*

$$f(x) \sim \sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx).$$

---

[3]) KACZMARZ – STEINHAUS, *Theorie der Orthogonalreihen* (Warszawa, 1935).

[4]) A. OSTROWSKI, Über Potenzreihen, die überkonvergente Abschnittsfolgen besitzen, *Sitzungsberichte der preußischen Akademie der Wissenschaften*, 1923, S. 185–192.

Es gäbe zwei Zahlen q und q' mit

$$1 < q' < q'^2 < q \tag{21}$$

und eine Folge der n-Werten

$$n_1 < n_2 < \ldots$$

mit

$$\frac{n_{k+1}}{n_k} > q \qquad k = 1, 2, \ldots \tag{22a}$$

und mit der Eigenschaft, daß $a_n = b_n = 0$ für alle n Werte, für welche

$$\frac{1}{q'} < \frac{n}{n_k} < q' \qquad n \neq n_k,\ k = 1, 2, \ldots \tag{22b}$$

gilt. Dann ist

$$\sum_{k=1}^{\infty} (|a_{n_k}| + |b_{n_k}|) < \infty.$$

B e w e i s. Es sei $x = x_0$ fest und $f(x)$ von oben beschränkt: $f(x) \leqq M$. Es sei wieder

$$\operatorname{sign}(a_{n_\nu} \cos n_\nu x_0 + b_{n_\nu} \sin n_\nu x_0) = \varepsilon_\nu, \qquad \nu = 1, 2, \ldots$$

und k die kleinste positive ganze Zahl so, daß

$$1 + \frac{1}{q^k - 1} < q', \quad 1 - \frac{1}{q^k - 1} > \frac{1}{q'}, \tag{23}$$

und

$$P_{mr}(y) = \prod_{i=1}^{m} (1 + \varepsilon_{ki+r} \cos n_{ki+r}\, y). \tag{24}$$

Ähnlich wie oben kann man einsehen, daß, wenn man $P_{mr}(y)$ als Cosinuspolynom aufschreibt, der Koeffizient von $\cos n_{ki+r}\, y$ gleich ε_{ki+r} ist und überhaupt nur solche Cosinusmultipla $\cos ny$ auftreten, für welche n in einem Intervall

$$\left[n_{ki+r}\left(1 - \frac{1}{q^k - 1}\right),\ n_{ki+r}\left(1 + \frac{1}{q^k - 1}\right) \right]$$

fällt. Nach (23) fallen also alle, in $P_{mr}(y)$ auftretenden Indizes a fortiori in die Intervalle

$$\left[n_{ki+r} \frac{1}{q'},\ n_{ki+r}\, q' \right], \qquad i = 0, 1, 2, \ldots, m.$$

Daraus folgt aber nach der Voraussetzung über die Fourier-Koeffizienten von $f(x)$, daß die Cosinusmultipla $\cos n_{k_{i+r}} y$ $(i=0, 1, 2, \ldots, m)$ die einzigen Multipla sind, welche in der Fourierentwickelung von $f(x)$ und in $P_{mr}(x)$ auftreten. Dann ist aber, wie oben

$$\sum_{i=0}^{m} |a_{n_{k_{i+r}}} \cos n_{k_{i+r}} x_0 + b_{n_{k_{i+r}}} \sin n_{k_{i+r}} x_0| = \frac{1}{\pi} \int_0^{2\pi} f(x) P_{mr}(x - x_0)\, dx,$$

woraus Satz III nach dem obigen Muster folgt.

Diesen Satz III kann man im Wesentlichen auf allgemeinere Orthogonalsysteme übertragen.

S a t z IV. *Es sei*

$$\varphi_0(x),\ \varphi_1(x),\ \varphi_2(x), \ldots$$

ein normiertes Orthogonalsystem (z. B. bezüglich des Intervalles $[0, 2\pi]$*) und es gelte*

$$(25) \qquad (0<)\, m \leqq |\varphi_n(x)| \leqq M, \qquad n = 0, 1, 2, \ldots$$

wo m *und* M *von* n *unabhängig sind.* (Solches System ist z. B. das Walsh-System.) *Dann existieren drei universelle Folgen von positiven Zahlen*

$$(26) \qquad \begin{array}{l} n_1' < n_2' < \ldots < n_k' < \ldots \\ n_1'' < n_2'' < \ldots < n_k'' < \ldots \\ n_1 < n_2 < \ldots < n_k < \ldots \end{array}$$

mit

$$(27a) \qquad n_1' < n_1'' < n_1 < \ldots < n_{k-1} < n_k' < n_k'' < n_k < \ldots$$

und

$$(27b) \qquad n_j'' - n_j' = N_j = \text{vorgeschrieben}$$

so daß, sobald in der Entwicklung

$$f(x) \sim \sum_{\nu=0}^{\infty} a_\nu \varphi_\nu(x)$$

einer einseitig beschränkten und L-integrablen $f(x)$ *alle* a_ν *mit*

$$n_j'' \leqq \nu \leqq n_{j+1}' \qquad \nu \neq n_j, \quad j = 1, 2, \ldots$$

verschwinden, dann gilt $\sum_{j=1}^{\infty} |a_{n_j}| < \infty$.

V o r b e m e r k u n g. In dem Spezialfalle $N_j = 1$ $(j = 1, 2, \ldots)$ hatte ich Satz IV in meiner Note „Über Orthogonalsysteme" (*Compositio Math.*, (1940), S. 372—375) ausgesprochen. Eine Lücke im Beweise habe ich in einem Nachtrag erfüllt, die aber im *Compositio Math.* nicht mehr erscheinen konnte. Mit Satz IV ist also auch diese Lücke ergänzt (gleichzeitig wird dabei der Satz verschärft, da ich in der ursprünglichen Fassung auch die stückweise Stetigkeit der $\varphi_\nu(x)$ voraussetzen mußte).

Es sei ohne Beschränkung der Allgemeinheit $M \geqq 1$, $m > \frac{2}{3}$, ferner

$$\varphi_n(x) \sim \sum_{l=0}^{\infty} (A_{nl} \cos lx + B_{nl} \sin lx). \tag{28}$$

Wir benötigen den folgenden Hilfssatz.

Hilfssatz A. *Aus der Folge $\varphi_0(x)$, $\varphi_1(x)$, ... kann man eine Teilfolge $\varphi_{n_1}(x)$, $\varphi_{n_2}(x)$ auswählen und zwei Zahlenfolgen $n'_1, n'_2, \ldots$; $n''_1, n''_2, \ldots$ angeben mit den folgenden Eigenschaften:*

a) $n'_1 < n''_1 < n_1 < \ldots < n_{k-1} < n'_k < n''_k < n_k < \ldots$; $n''_k - n'_k = N_k$

b) *zu jeder $\varphi_{n_k}(x)$ existiert ein trigonometrisches Polynom*

$$F_k(x) = \sum_{l=i_k}^{I_k} (c_{kl} \cos lx + d_{kl} \sin lx), \tag{29}$$

für welche

$$|\varphi_{n_k}(x) - F_k(x)| < \frac{1}{2^k}$$

überall mit der Ausnahme einer Menge E_k vom Maße $\leqq \frac{1}{4^k}$ gilt; in den Punkten von E_k gilt

$$|F_k(x)| \leqq 2M; \tag{30}$$

c) *es gilt $I_1 > i_1$ und für $k \geqq 2$*

$$I_k > i_k > 2 \sum_{\nu=1}^{k-1} I_\nu;$$

d) *für jedes $n > n'_{k+1}$ gilt*

$$\sum_{l=0}^{\left[\frac{3}{2} I_k\right]} (|A_{ln}| + |B_{ln}|) < \frac{1}{3^k N_k}, \qquad k = 1, 2, \ldots;$$

e) *für jedes $n \leqq n''_k$ gilt*

$$\sum_{l=\left[\frac{1}{2} i_k\right]}^{\infty} (A_{ln}^2 + B_{ln}^2) < \frac{1}{4^k} \min_{1 \leqq \mu \leqq k} \frac{1}{N_\mu^2}.$$

Bemerkung. Von diesen Zahlen n_k, n'_k, n''_k werden wir beweisen, daß sie die im Satz IV behaupteten Eigenschaften besitzen.

Beweis des Hilfssatzes. Es sei $n_1'=0$, $n_1''=N_1$. Da wegen der Normierung des Systems $\varphi_\nu(x)$ und (28)

$$\sum_{l=0}^{\infty}(A_{l\,n}^2+B_{l\,n}^2)=1 \tag{31}$$

gilt, so gibt es eine kleinste ganze Zahl i_1 so, daß für $n_1'\leqq n\leqq n_1''$

$$\sum_{l=\left[\frac{1}{2}i_1\right]}^{\infty}(A_{l\,n}^2+B_{l\,n}^2)\leqq\frac{1}{4^1}\,\frac{1}{N_1^2}. \tag{32}$$

Da bekanntlich bei festem ν

$$\lim_{r\to\infty}\int_0^{2\pi}\varphi_r(x)\cos\nu x\,dx=\lim_{r\to\infty}\int_0^{2\pi}\varphi_r(x)\sin\nu x\,dx=0 \tag{33}$$

gilt, gibt es ein kleinstes Index n_1, für welches

$$n_1>n_1'' \tag{34a}$$

und für jedes $n\geqq n_1$

$$\sum_{l=0}^{i_1-1}(|A_{l\,n}|+|B_{l\,n}|)\leqq\frac{1}{10}. \tag{34b}$$

Mit diesem n_1 bilden wir $\varphi_{n_1}(x)$; dann gibt es bekanntlich ein kleinstes I_1 so, daß

$$I_1>i_1 \tag{35a}$$

und, wenn $\sigma_{I_1}(x)$ die I_1-te Cesàro-Mittel erster Ordnung der Fourierreihe von $\varphi_{n_1}(x)$ bedeutet,

$$\int_0^{2\pi}(\varphi_{n_1}(x)-\sigma_{I_1}(x))^2\,dx<\frac{1}{100}. \tag{35b}$$

Dann ist aber

$$|\varphi_{n_1}(x)-\sigma_{I_1}(x)|\leqq\frac{4}{10} \tag{36}$$

mit Ausnahme einer Menge E_1 vom Maße $<\frac{1}{4}$. Da nach dem bekannten Fejérschen Satze $|\sigma_{I_1}(x)|\leqq M$ *überall* gilt, so hat man, wenn $\sigma_{I_1,i_1}(x)$ die (i_1-1)-te Partialsumme von $\sigma_{I_1}(x)$ bedeutet, für die Funktion

$$F_1(x)=\sigma_{I_1}(x)-\sigma_{I_1,i_1}(x)$$

auf E_1 wegen (34b) und $M\geqq 1$ offenbar

$$|F_1(x)|\leqq M+\frac{1}{10}<2M \tag{37a}$$

und auf $\bar{E}_1$ wegen (36) und (34b)

$$(37\text{b}) \qquad |F_1(x)-\varphi_{n_1}(x)| \leqq |\varphi_{n_1}(x)-\sigma_{I_1}(x)|+|\sigma_{I_1,\,i_1}(x)| \leqq \frac{1}{2^1}.$$

Dann existiert aber nach (33) eine kleinste ganze Zahl n_2' so, daß für jedes $n \geqq n_1'$

$$(38) \qquad \sum_{l=0}^{\left[\frac{3}{2} I_1\right]} (|A_{l_n}|+|B_{l_n}|) < \frac{1}{3^1 N_1}$$

gilt. So können wir voraussetzen daß

$$\begin{array}{llll}
n_1', & n_2' & ,\ldots, & n_k',\ n_{k+1}' \\
n_1'', & n_2'' & ,\ldots, & n_k'' \\
i_1, & i_2 & ,\ldots, & i_k \\
n_1, & n_2 & ,\ldots, & n_k \\
I_1, & I_2 & ,\ldots, & I_k \\
F_1(x), & F_2(x), & \ldots, & F_k(x) \\
E_1, & E_2 & ,\ldots, & E_k
\end{array}$$

bereits definiert sind; wir zeigen, daß das obige sukzessive Konstruktionsverfahren nicht abbricht. n_{k+1}'' sei natürlich gleich $n_{k+1}'+N_{k+1}$. Dann sei i_{k+1} als die kleinste ganze Zahl definiert, für welche

$$(39\text{a}) \qquad i_{k+1} > 2\sum_{\nu=1}^{k} I_\nu$$

und für welche bei jedem $n \leqq n_{k+1}''$

$$(39\text{b}) \qquad \sum_{l=\left[\frac{1}{2} i_{k+1}\right]}^{\infty} (A_{l_n}^2+B_{l_n}^2) < \frac{1}{4^{k+1}} \min_{1\leqq j\leqq k+1} \frac{1}{N_j^2}$$

(das geht wegen (31)). Nach (33) gibt es ein kleinstes Index n_{k+1} so, daß

$$(40\text{a}) \qquad n_{k+1} > n_{k+1}''$$

und für jedes $n \geqq n_{k+1}$

$$(40\text{b}) \qquad \sum_{l=0}^{i_{k+1}-1} (|A_{l_n}|+|B_{l_n}|) \leqq \frac{1}{10^{k+1}}.$$

Mit diesem n_{k+1} bilden wir $\varphi_{n_{k+1}}(x)$; dann gibt es ein kleinstes ganzes I_{k+1} so, daß

$$(41\text{a}) \qquad I_{k+1} > i_{k+1}$$

und

$$(41\text{b}) \qquad \int_0^{2\pi} (\varphi_{n_{k+1}}(x) - \sigma_{I_{k+1}}(x))^2 \, dx < \frac{1}{10^{2k+2}}.$$

Dann gilt

$$(42) \qquad |\varphi_{n_{k+1}}(x) - \sigma_{I_{k+1}}(x)| < \left(\frac{4}{10}\right)^{k+1}$$

mit Ausnahme einer Menge E_{k+1} vom Maße $< \frac{1}{4^{k+1}}$. Dann ist aber, wenn wir mit den obigen Bezeichnungen

$$F_{k+1}(x) = \sigma_{I_{k+1}}(x) - \sigma_{I_{k+1},\, i_{k+1}}(x)$$

wählen, wie oben, auf E_{k+1}

$$(43\text{a}) \qquad |F_{k+1}(x)| \leqq 2M$$

und auf $\overline{E}_{k+1}$

$$(43\text{b}) \qquad |F_{k+1}(x) - \varphi_{n_{k+1}}(x)| \leqq \frac{1}{2^{k+1}}.$$

Mit diesen I_{k+1} existiert nach (33) ein kleinstes ganzes n'_{k+2} so, daß

$$n'_{k+2} > n_{k+1}$$

und für jedes $n \geqq n'_{k+1}$

$$\sum_{l=0}^{\left[\frac{3}{2} I_{k+1}\right]} (|A_{l n}| + |B_{l n}|) < \frac{1}{3^{k+1} N_{k+1}}.$$

Somit ist der Hilfssatz bewiesen.

Beweis des Satzes IV. Es sei mit den obigen n_k

$$(44\text{a}) \qquad f(x) \sim \sum_{\nu=1}^{n_L} a_\nu \varphi_\nu(x),$$

$$(44\text{b}) \qquad f(x) \leqq K.$$

Über $f(x)$ wird also vorausgesetzt, daß mit den obigen n'_k und n''_k

$$(45) \qquad a_\nu = 0 \quad \text{für} \quad n''_j \leqq \nu \leqq n'_{j+1}, \quad \nu \neq n_j, \; j = 1, 2, \ldots, L.$$

Es sei mit den obigen Polynomen $F_k(x)$

$$(46) \qquad P_d(x) = \prod_{j=1}^{d} \left(1 + \operatorname{sign} a_{n_j} \frac{F_j(x)}{4M}\right), \qquad d = 1, 2, \ldots, L.$$

Nach (30) ist $P_d(x)$ nichtnegativ; ferner gilt nach c)

$$(47) \qquad \int_0^{2\pi} |P_d(x)| \, dx = \int_0^{2\pi} P_d(x) \, dx = 2\pi \qquad (d = 1, 2, \ldots, L).$$

Da offenbar

$$(48)\quad P_d(x)=P_{d-1}(x)+\frac{1}{4M}\operatorname{sign}a_{n_d}F_d(x)\,P_{d-1}(x)\qquad (d=2,\ldots,L)$$

gilt, so folgt

$$\int_0^{2\pi}P_d(x)F_d(x)\,dx=\int_0^{2\pi}P_{d-1}(x)F_d(x)\,dx+\frac{1}{4M}\operatorname{sign}a_{n_d}\int_0^{2\pi}P_{d-1}(x)\,F_d^2(x)\,dx;$$

da ferner nach b)

$$\int_{\bar{E}_d}P_{d-1}(x)\,dx=\int_0^{2\pi}P_{d-1}(x)\,dx-\int_{E_d'}P_{d-1}(x)\,dx>2\pi-\frac{1}{4^d}\left(\frac{3}{2}\right)^{d+1}>\frac{3\pi}{2},$$

so gilt nach c) und (47)

$$(49)\quad \operatorname{sign}a_{n_d}\int_0^{2\pi}P_d(x)\,F_d(x)\,dx=\frac{1}{4M}\int_0^{2\pi}P_{d-1}(x)\,F_d^2\,(x)\,dx>$$

$$>\frac{1}{4M}\int_{\bar{E}_d}P_{d-1}(x)F_d^2(x)\,dx>\frac{\left(m-\frac{1}{2^d}\right)^2}{4M}\int_{\bar{E}_d}P_{d-1}(x)\,dx>\frac{1}{144M}\,\frac{3\pi}{2}.$$

Wenn $j<k$, so ist nach (48)

$$(50)\quad \int_0^{2\pi}P_k(x)\,F_j(x)\,dx=$$

$$=\int_0^{2\pi}P_{k-1}(x)\,F_j(x)\,dx+\frac{1}{4M}\operatorname{sign}a_{n_k}\int_0^{2\pi}F_k(x)\,P_{k-1}(x)\,F_j(x)\,dx.$$

Da aber $P_{k-1}(x)$ ein trigonometrisches Polynom von der Ordnung $l_1+l_2+\ldots+l_{k-1}$, $F_j(x)$ ein ebensolches von der Ordnung

$$l_j\leqq l_{k-1}\leqq l_1+l_2+\ldots+l_{k-1}$$

bedeuten, ist die Ordnung von $P_{k-1}(x)\,F_j(x)$ kleiner als $2\sum_{\nu=1}^{k-1}l_\nu<i_k$ (siehe c)), also das zweite Glied rechts in (50) verschwindet nud es gilt

$$\int_0^{2\pi}P_k(x)\,F_j(x)\,dx=\int_0^{2\pi}P_{k-1}(x)\,F_j(x)\,dx,\qquad j<k,$$

daher

$$(51)\quad \int_0^{2\pi}P_k(x)\,F_j(x)\,dx=\int_0^{2\pi}P_j(x)\,F_j(x)\,dx,\qquad j<k$$

und endlich nach (49) mit $d=j$

$$(52)\qquad \operatorname{sign} a_{n_j} \int_0^{2\pi} P_k(x) F_j(x)\, dx > \frac{1}{144M}\,\frac{3\pi}{2}.$$

Wir benötigen aber eine Abschätzung der Integrale $\int_0^{2\pi} P_k(x)\varphi_{n_j}(x)dx$. Es gilt wegen (29), (30), (47)

$$\left|\int_0^{2\pi} P_j(x)\,\varphi_{n_j}(x)\,dx - \int_0^{2\pi} P_j(x)\,F_j(x)\,dx\right| =$$

$$= \left|\int_0^{2\pi} P_j(x)\,(\varphi_{n_j}(x) - F_j(x))\,dx\right| < \frac{1}{2^j}\,2\pi + 2M\int_{E_j} |P_j(x)|\,dx <$$

$$< \frac{2\pi}{2^j} + 2M\frac{1}{4^j}\left(\frac{3}{2}\right)^j < \frac{10M}{2^j},$$

also nach (49)

$$(53)\qquad \operatorname{sign} a_{n_j} \int_0^{2\pi} P_j(x)\,\varphi_{n_j}(x)\,dx > \frac{1}{44M}\,\frac{3\pi}{2} - \frac{10M}{2^j}.$$

Ferner gilt für $1 \leqq j < d \leqq L$ nach (48)

$$(54)\qquad \int_0^{2\pi} (P_d(x) - P_{d-1}(x))\,\varphi_{n_j}(x)\,dx =$$

$$= \frac{1}{4M}\operatorname{sign} a_{n_d} \int_0^{2\pi} P_{d-1}(x)\,F_d(x)\,\varphi_{n_j}(x)\,dx.$$

Nach (28), (54) gilt wegen

$$F_d(x)\,P_{d-1}(x) = c_0 \cos\left(i_d - (l_1 + l_2 + \ldots + l_{d-1})\right)x + \ldots$$

und wegen c)

$$\left|\int_0^{2\pi} (P_d(x) - P_{d-1}(x))\,\varphi_{n_j}(x)\,dx\right| =$$

$$(55)\qquad = \frac{1}{4M}\int_0^{2\pi} P_{d-1}(x)\,F_d(x)\left[\sum_{l=\left[\frac{1}{2} i_d\right]}^{\infty} (A_{l n_j}\cos lx + B_{l n_j}\sin lx)\right] dx <$$

$$< \frac{1}{4M}\sqrt{\sum_{l=\left[\frac{1}{2} i_d\right]}^{\infty} (A_{l n_j}^2 + B_{l n_j}^2)}\;\sqrt{\int_0^{2\pi} P_{d-1}^2(x)\,F_d^2(x)\,dx}.$$

Der erste Faktor rechts in (55) ist nach $j < d$ und e) kleiner als

$$(56a)\qquad \frac{1}{2^d}\min_{1 \leqq \mu \leqq d} \frac{1}{N_\mu},$$

der zweite wegen (47) kleiner als

$$\text{(56b)}\quad 2M\left(\int_0^{2\pi} P_{d-1}^2(x)\,dx\right)^{1/2} <$$

$$< 2M \max_{0\leqq x\leqq 2\pi} |P_{d-1}(x)|\sqrt{2\pi} < 20M\left(\frac{3}{2}\right)^d,$$

also folgt aus (55), (56a), (56b)

$$\text{(57)}\quad \left|\int_0^{2\pi} P_d(x)-P_{d-1}(x))\,\varphi_{n_j}(x)\,dx\right| < 5\left(\frac{3}{2}\right)^d \min_{1\leqq\mu\leqq d}\frac{1}{N_\mu}.$$

Wenn wir (57) mit $d=k,\ k-1,\ldots,j+1$ anwenden und summieren, gewinnen wir

$$\left|\int_0^{2\pi} P_k(x)\,\varphi_{n_j}(x)\,dx - \int_0^{2\pi} P_j(x)\,\varphi_{n_j}(x)\,dx\right| < 20\left(\frac{3}{4}\right)^{j+1}\min_{1\leqq\mu\leqq j+1}\frac{1}{N_\mu},$$

also nach (53) für $j<k$

$$\text{(58)}\quad \operatorname{sign} a_{n_j}\int_0^{2\pi} P_k(x)\,\varphi_{n_j}(x)\,dx >$$

$$> \frac{1}{144M}\,\frac{3\pi}{2} - \frac{10M}{2^j} - 20\left(\frac{3}{4}\right)^{j+1}\min_{1\leqq\mu\leqq j+1}\frac{1}{N_\mu}.$$

Wir benötigen noch eine Abschätzung der Integrale $\int_0^{2\pi} P_j(x)\,\varphi_n(x)\,dx$, wenn $n>n'_{j+1}$. Da die Ordnung von $P_j(x)$ nach c) $(l_1+l_2+\ldots+l_{j-1})+l_j<\frac{l_j}{2}+l_j<\frac{3}{2}l_j$ ist, gilt wegen (47) und d)

$$\int_0^{2\pi} P_j(x)\,\varphi_n(x)\,dx = \int_0^{2\pi} P_j(x)\left[\sum_{l=0}^{\left[\frac{3}{2}l_j\right]}(A_{l_n}\cos lx+B_{l_n}\sin lx)\right]dx <$$

$$\text{(59)}\qquad < \sum_{l=0}^{\left[\frac{3}{2}l_j\right]}(|A_{l_n}|+|B_{l_n}|)\int_0^{2\pi}|P_j(x)|\,dx =$$

$$= 2\pi\sum_{l=0}^{\left[\frac{3}{2}l_j\right]}(|A_{l_n}|+|B_{l_n}|) < \frac{2\pi}{3^j N_j}.$$

Wenn ferner $j \leqq d \leqq k$ und $n'_j \leqq n \leqq n''_j$, dann ist wegen (48), wie in (55)

$$\left|\int_0^{2\pi} (P_d(x)-P_{d-1}(x))\,\varphi_n(x)\,dx\right| = \frac{1}{4M}\left|\int_0^{2\pi} P_{d-1}(x)\,F_d(x)\,\varphi_n(x)\,dx\right| <$$

$$< \frac{1}{4M}\sqrt{\sum_{l=\left[\frac{1}{2}i_d\right]}^{\infty} (A_{ln}^2+B_{ln}^2)}\sqrt{\int_0^{2\pi} P_{d-1}^2(x)\,F_d^2(x)\,dx} <$$

$$< \frac{1}{2^d}\min_{1\leqq\mu\leqq d}\frac{1}{N_\mu}10\left(\frac{3}{2}\right)^{\frac{d}{2}},$$

also erhalten wir durch Summation für $d=j, j+1, \ldots, k$

$$\left|\int_0^{2\pi} P_k(x)\,\varphi_n(x)\,dx - \int_0^{2\pi} P_{j-1}(x)\,\varphi_n(x)\,dx\right| < 50\min_{1\leqq\mu\leqq j}\frac{1}{N_\mu}\left(\frac{3}{8}\right)^{\frac{j}{2}}.$$

Nach (59) gilt also für $n'_j \leqq n \leqq n''_j$ $(j \leqq k)$

$$\left|\int_0^{2\pi} P_k(x)\,\varphi_n(x)\,dx\right| < \frac{2\pi}{3^{j-1}N_{j-1}} + 50\left(\frac{3}{8}\right)^{\frac{j}{2}}\min_{1\leqq\mu\leqq j}\frac{1}{N_\mu} < \tag{60}$$

$$< 100\left(\frac{3}{8}\right)^{\frac{j}{2}}\min_{1\leqq\mu\leqq j}\frac{1}{N_\mu}.$$

Es sei endlich

$$I = \int_0^{2\pi} f(x)\,P_L(x)\,dx. \tag{61}$$

Es ist wegen (47) und (44b)

$$I < 2\pi K. \tag{62a}$$

Ferner folgt nach (44a) und (45)

$$I = \sum_{n=1}^{n_L} a_n \int_0^{2\pi} P_L(x)\,\varphi_n(x)\,dx = \sum_{j=1}^{L}\sum_{n=n'_j}^{n''_j} a_n \int_0^{2\pi} P_L(x)\,\varphi_n(x)\,dx +$$

$$+ \sum_{j=1}^{L} a_{n_j}\int_0^{2\pi} P_L(x)\,\varphi_{n_j}(x)\,dx,$$

also nach (58) und (60)

$$I > \sum_{j=1}^{L} |a_{n_j}| \left(\frac{1}{144M} \frac{3\pi}{2} - \frac{10M}{2^j} - 20 \left(\frac{3}{4}\right)^{j+1} \frac{1}{\max\limits_{1 \leq \mu \leq j+1} N_\mu} \right) -$$

$$(62b) \qquad - \sum_{j=1}^{L} 100 \left(\frac{3}{8}\right)^{\frac{j}{2}} N_j \min_{1 \leq \mu \leq j} \frac{1}{N_\mu} \max_{n_j' \leq \nu \leq n_j''} |a_\nu| >$$

$$> \sum_{j=1}^{L} |a_{n_j}| \left(\frac{1}{144M} \frac{3\pi}{2} - \frac{10M}{2^j} - 20 \left(\frac{3}{4}\right)^{j+1} \min_{1 \leq \mu \leq j+1} \frac{1}{N_\mu} \right) -$$

$$- 400 \int_0^{2\pi} |f(x)| dx.$$

Aus (62a) und (62b) folgt unmittelbar Satz IV.

§. 2.

Es sei $\varphi_0(x)$, $\varphi_1(x), \ldots$ ein beliebiges normiertes Orthogonalsystem bezüglich $[0, 2\pi]$; $\varepsilon_1, \varepsilon_2, \ldots$ eine Zahlenfolge und $n_1 < n_2 < \ldots$ eine Indexfolge. Die Probleme, welche in diesem Paragraphen behandelt werden, sind folgender Art. Es sollen hinreichende Bedingungen über die ε_ν und n_ν angegeben werden, welche die Lösbarkeit des Gleichungsystems

$$(63) \qquad \int_0^{2\pi} f(x)\, \varphi_{n_k}(x)\, dx = \varepsilon_k$$

in einer vorgegebene Funktionenklasse sichern. Wenn das Orthogonalsystem die Bedingung (25) erfüllt und jede der Funktionen $\varphi_\nu(x)$ stückweise stetig ist, hatte ich[5]) die Existenz einer Indexfolge bewiesen, bei welchen das System sicherlich L-integrable Lösung besitzt, falls nur die ε_ν eine Nullfolge bilden. Den Beweis hatte ich auf einen Absolutkonvergenz-Satz und auf einen Satz von S. Banach[6]) gegründet. Eine Lücke dieses Beweises habe ich im § 1. ausgefüllt und gleichzeitig den Satz verschärft. So ergibt sich also der

Satz V. *Es sei* $\varphi_0(x), \varphi_1(x), \ldots$ *ein normiertes Orthogonalsystem mit der Eigenschaft* (25). *Dann existiert eine Indexfolge* $n_1 < n_2 < \ldots$ *so, daß das System* (63) *eine L-integrable Lösung besitzt, falls nur die Folge* ε_ν *eine Nullfolge bildet.*

---

5) S. Sidon, Über Orthogonalsysteme, *Compositio Math.*, 7 (1940), S. 372—375.

6) S. Banach, Über einige Eigenschaften der lakunären trigonometrischen Reihen, *Studia Math.*, 2 (1930), S. 207—228.

Wie ich nach der Abfassung meiner Compositio-Note [5]) erkannte, bewies Herr J. MARCINKIEWICZ in seiner Arbeit[7]) u. a. den folgenden Satz: wenn die Funktionen $\varphi_0(x), \varphi_1(x), \dots$ im $[0, 2\pi]$ überall stetig sind und $\overline{\lim\limits_{n\to\infty}} \int\limits_0^{2\pi} |\varphi_n(x)|\,dx > 0$, so hat das Gleichungsystem (63) eine L-integrable Lösung, falls nur ε_ν eine Nullfolge bildet. Satz V scheint sich mit der von MARCINKIEWICZ angewandten Beweismethode nicht zu ergeben.

Wenn $\varphi_0(x), \varphi_1(x), \dots$ das trigonometrische System bedeutet, bewies ich[8]), daß das System

$$(64)\qquad \frac{1}{\pi}\int\limits_0^{2\pi} f(x)\cos n_k x\,dx = \varepsilon_k', \quad \frac{1}{\pi}\int\limits_0^{2\pi} f(x)\sin n_k x\,dx = \varepsilon_k''$$

sicherlich lösbar ist in der Klasse C der überall stetigen Funktionen, wenn die Indexfolge n_k eine B_2^*-Folge bildet, falls nur $\sum\limits_{k=1}^{\infty} (\varepsilon_k'^2 + \varepsilon_k''^2) < \infty$. B_2^*-Folge nennt man eine Folge $n_1, n_2, \dots$, wenn die Koeffizientenfolge der formalen Laurentreihe

$$\left(\sum_{k=1}^{\infty} z^{-n_k} + \sum_{k=1}^{\infty} z^{n_k}\right)^2 = \sum_{\nu=-\infty}^{+\infty} A_\nu z^\nu$$

gleichmäßig beschränkt ist. Solche Indexfolgen können viel „dichter" sein als die lakunären Folgen; eine B_2^*-Folge kann die Abschätzung $n_k = O(k^3)$ erfüllen. Ich bewies nur die Existenz einer solchen Folge und ich weiß nicht, ob es nicht B_2^*-Folgen mit $n_k = O(k^{2+\varepsilon})$ existieren. In der Richtung meines Satzes [8]) bewies ich

Satz VI. *Es seien $\varepsilon_k', \varepsilon_k''$ zwei Zahlenfolgen mit der Eigenschaft, daß die Reihe $\sum\limits_{\nu=1}^{\infty} \nu^{\varepsilon}(\varepsilon_k'^2 + \varepsilon_k''^2)$ bei jedem festen positiven ε konvergiert. Dann ist des System*

---

[7]) J. MARCINKIEWICZ, Sur les séries orthogonales, *Studia Math.*, **8** (1939), S. 1–27.

[8]) S. SIDON, Ein Satz über trigonometrische Polynome und seine Anwendung in der Theorie der Fourier-Reihen, *Math. Annalen*, **106** (1932), S. 536–539. In meiner Arbeit: Bemerkungen über Fourier und Potenreihen, *diese Acta*, 7 (1936), S. 85–94 bewies ich, daß $\left(\sum\limits_{k=1}^{\infty} z^{n_k}\right)^2 = \sum\limits_{k=1}^{\infty} B_k z^k$ $B_k = O(1)$ genügt.

$$(65) \qquad \frac{1}{\pi}\int_0^{2\pi} f(x)\cos k^2 x\,dx = \varepsilon_k', \quad \frac{1}{\pi}\int_0^{2\pi} f(x)\sin k^2 x\,dx = \varepsilon_k''$$

sicherlich lösbar in der Klasse C.

Der wahre Grund dieses Satzes liegt in der zahlentheoretischen Tatsache, daß die Lösungzahl der Gleichung $x^2+y^2=n$ — wie bekannt — $O(n^\varepsilon)$ ist. Mit gleicher Mühe beweisen wir gleich den allgemeinen

Satz VII. *Für die Indexfolge $n_1 < n_2 < \ldots$ gelte*

$$(66a) \qquad \frac{n_{2k}}{n_k} > q > 1 \qquad k = 1, 2, \ldots$$

und für die Lösungzahl λ_n der Gleichung $n_k + n_l = n$ gelte die Abschätzung

$$(66b) \qquad \lambda_n = O(\varphi(n)),$$

wo $\varphi(x)$ eine positive, nichtabnehmende Funktion mit der Eigenschaft

$$(66c) \qquad \varphi(2x) < c_1\varphi(x)$$

bedeutet (c_1 absolute Konstante). Es sei ferner $\beta_1 < \beta_2 < \ldots$ eine Zahlenfolge, für welche

$$(66d) \qquad \overline{\lim_{k\to\infty}} \frac{\varphi(n_{2k})}{\beta_k} < \infty$$

und

$$(66e) \qquad \sum_{k=1}^{\infty} \varphi(n_{2k})\left(\frac{1}{\beta_k} - \frac{1}{\beta_{k+1}}\right) < \infty.$$

Wenn die Reihe $\sum_{\nu=1}^{\infty} \beta_\nu(\varepsilon_\nu'^2 + \varepsilon_\nu''^2)$ konvergiert, so ist das System (64) in Klasse C sicherlich lösbar[9]*.*

Beweis. Es sei $z = re^{i\vartheta}$, $b_0 = 0$, a_ν, b_ν reell und

$$(67) \qquad f(z) = \sum_{\nu=0}^{n} (a_\nu - ib_\nu) z^\nu, \quad \Re f(z) = \sum_{\nu=0}^{n} (a_\nu \cos\nu\vartheta + b_\nu \sin\nu\vartheta).$$

Für eine positive Funktion $\psi(\vartheta)$ gilt nach der Hölderschen Ungleichung

---

[9]) Die Konvergenz der Reihe $\sum_{\nu=1}^{\infty}(\varepsilon_\nu'^2 + \varepsilon_\nu''^2)$ ist offenbar notwendig, damit $f(x) \in C$ sei. Für $\varphi(n) = c$, $\beta_\nu = c$ ergibt sich also eine Verschärfung meines Satzes in [8]).

$$\int_0^{2\pi} \psi^2(\vartheta)\,d\vartheta = \int_0^{2\pi} \psi(\vartheta)^{\frac{2}{3}}\,\psi(\vartheta)^{\frac{4}{3}}\,d\vartheta \leqq \Big(\int_0^{2\pi} \psi(\vartheta)\,d\vartheta\Big)^{\frac{2}{3}} \Big(\int_0^{2\pi} \psi(\vartheta)^4\,d\vartheta\Big)^{\frac{1}{3}},$$

d. h.

$$\frac{\Big(\int_0^{2\pi} \psi(\vartheta)\,d\vartheta\Big)^2}{\int_0^{2\pi} \psi^2(\vartheta)\,d\vartheta} < \frac{\Big(\int_0^{2\pi} \psi^2(\vartheta)\,d\vartheta\Big)^2}{\int_0^{2\pi} \psi^4(\vartheta)\,d\vartheta}.$$

Es sei hier $\psi(\vartheta) = |\Re f(z)|$; dann gilt

$$\frac{1}{2\pi} \int_{|z|=1} |f(z)|^2\,d\vartheta = a_0^2 + \sum_{\nu=1}^{n} (a_\nu^2 + b_\nu^2) \leqq$$

$$\leqq 2\Big[a_0^2 + \frac{1}{2}\sum_{\nu=1}^{n} (a_\nu^2 + b_\nu^2)\Big] = 2\,\frac{1}{2\pi} \int_0^{2\pi} (\Re f(z))^2\,d\vartheta,$$

also

$$\frac{\Big(\int_0^{2\pi} |\Re f(z)|\,d\vartheta\Big)^2}{\int_0^{2\pi} (\Re f(z))^2\,d\vartheta} > \frac{\Big(\frac{1}{2}\int_0^{2\pi} |f(z)|^2\,d\vartheta\Big)^2}{\int_0^{2\pi} (\Re f(z))^4\,d\vartheta}.$$

Nach einer Ungleichung von M. Riesz gilt

$$\int_0^{2\pi} (\Re f(z))^4\,d\vartheta < c_2 \int_0^{2\pi} |f(z)|^4\,d\vartheta$$

mit einer absoluten Konstanten c_2; also

$$\frac{\Big(\int_0^{2\pi} |\Re f(z)|\,d\vartheta\Big)^2}{\int_0^{2\pi} (\Re f(z))^2\,d\vartheta} > \frac{1}{4c_2}\,\frac{\Big(\int_0^{2\pi} |f(z)|^2\,d\vartheta\Big)^2}{\int_0^{2\pi} |f(z)|^4\,d\vartheta}. \tag{68}$$

Wenn $f(z) = \sum_{\nu=1}^{k} \alpha_\nu z^{n_\nu}$, $\alpha_\nu = a_\nu - i b_\nu$, so gilt

$$f^2(z) = \sum_{\nu \leqq 2n_k} g_\nu z^\nu,$$

wo

$$g_\nu = \sum_{n_j + n_l = \nu} \alpha_j \alpha_l.$$

Nach der Voraussetzung ist die Anzahl der Glieder in der rechtseitigen Summe $\leqq \varphi(\nu)$; also gilt nach der Cauchyschen Ungleichung

$$|g_\nu|^2 \leqq \varphi(\nu) \cdot \sum_{n_j + n_l = \nu} |\alpha_j|^2 |\alpha_l|^2$$

und wegen (66c)

$$\frac{1}{2\pi} \int_0^{2\pi} |f(z)|^4 \, d\vartheta = \sum_{\nu \leqq 2n_k} |g_\nu|^2 \leqq$$

$$\leqq \varphi(2n_k) \left(\sum_{j=1}^{k} |\alpha_j|^2 \right)^2 \leqq c_1 \varphi(n_k) \left(\frac{1}{2\pi} \int_0^{2\pi} |f(z)|^2 \, d\vartheta \right)^2.$$

Aus dieser Ungleichung und (68) folgt bei jeder Wahl von a_ν und b_ν und k

$$(69) \quad \left(\int_0^{2\pi} \left| \sum_{\nu=1}^{k} (a_\nu \cos n_\nu x + b_\nu \sin n_\nu x) \right| dx \right)^2 >$$

$$> \frac{c_3}{\varphi(n_k)} \int_0^{2\pi} \left[\sum_{\nu=1}^{k} a_\nu \cos n_\nu x + b_\nu \sin n_\nu x \right]^2 dx$$

mit absolut-konstantem c_3. Es sei

$$(70) \qquad \sum_{\nu=1}^{K} (a_\nu \cos n_\nu x + b_\nu \sin n_\nu x) = F(x)$$

und $S_n(x)$ bedeute das n-te Cesàro-Mittel erster Ordung von $F(x)$ (wo natürlich bei der Mittelbildung auch die verschwindenden Glieder zu rechnen sind). Nach (66a) gilt dann für $j \leqq K$

$$\sum_{h=1}^{j} (a_h^2 + b_h^2) \leqq \left(\frac{q}{q-1} \right)^2 \sum_{h=1}^{j} \left(\frac{n_{2j} - n_h + 1}{n_{2j} + 1} \right)^2 (a_h^2 + b_h^2) =$$

$$= \left(\frac{q}{q-1} \right)^2 \frac{1}{2\pi} \int_0^{2\pi} S_{n_{2j}}^2(x) \, dx,$$

also nach (69) für $1 \leqq j \leqq K$

$$(71) \quad \sum_{h=1}^{j} (a_h^2 + b_h^2) < c_4(q) \varphi(n_{2j}) \left(\int_0^{2\pi} |S_{n_{2j}}(x)| \, dx \right)^2 <$$

$$< c_4(q) \varphi(n_{2j}) \left(\int_0^{2\pi} |F(x)| \, dx \right)^2$$

wegen einer bekannten Ungleichung. Es sei $s_j = \sum_{h=1}^{j} (a_h^2 + b_h^2)$; aus (69) folgt, wenn wir $a_\nu = a'_\nu \sqrt{\beta_\nu}$, $b_\nu = b'_\nu \sqrt{\beta_\nu}$ setzen,

$$U = \sum_{h=1}^{K} (a_h'^2 + b_h'^2) =$$

$$(72) \quad = \sum_{h=1}^{K} (a_h^2 + b_h^2) \frac{1}{\beta_h} = s_1 \frac{1}{\beta_1} + (s_2 - s_1) \frac{1}{\beta_2} + \dots + (s_K - s_{K-1}) \frac{1}{\beta_K} =$$

$$= s_1 \left(\frac{1}{\beta_1} - \frac{1}{\beta_2} \right) + \dots + s_{K-1} \left(\frac{1}{\beta_{K-1}} - \frac{1}{\beta_K} \right) + s_K \frac{1}{\beta_K}$$

d. h. aus (71) und (72)

$$U < c_4(q) \left(\int_0^{2\pi} \left| \sum_{\nu=1}^{K} \sqrt{\beta_\nu} (a'_\nu \cos n_\nu x + b'_\nu \sin n_\nu x) \right| dx \right)^2 \cdot$$

$$\cdot \left[\varphi(n_2) \left(\frac{1}{\beta_1} - \frac{1}{\beta_2} \right) + \dots + \varphi(n_{2K-2}) \left(\frac{1}{\beta_{K-1}} - \frac{1}{\beta_K} \right) + \frac{\varphi(n_{2K})}{\beta_K} \right] <$$

$$< c_5(q) \left(\int_0^{2\pi} \sum_{\nu=1}^{K} \sqrt{\beta_\nu} (a'_\nu \cos n_\nu x + b'_\nu \sin n_\nu x) \right| dx \right)^2$$

nach (66d) und (66e). Es gilt also die Ungleichung

$$(73) \quad \int_0^{2\pi} \left| \sum_{\nu=1}^{K} \sqrt{\beta_\nu} (a'_\nu \cos n_\nu x + b'_\nu \sin n_\nu x) \right| dx >$$

$$> c_6(q) \left[\int_0^{2\pi} \left(\sum_{\nu=1}^{K} (a'_\nu \cos n_\nu x + b'_\nu \sin n_\nu x) \right)^2 dx \right]^{1/2}.$$

Wie ich bewies[8], ist der Satz VII im Spezialfalle $\varphi(n) = c$ eine Folge der Ungleichung

$$(73a) \quad \int_0^{2\pi} \left| \sum_{\nu=1}^{K} (a'_\nu \cos n_\nu x + b'_\nu \sin n_\nu x) \right| dx >$$

$$> c_7(q) \left[\int_0^{2\pi} \left(\sum_{\nu=1}^{K} (a'_\nu \cos n_\nu x + b'_\nu \sin n_\nu x) \right)^2 dx \right]^{1/2}.$$

Der allgemeine Fall ist ganz analogerweise eine Folge von (73).

In der Arbeit [8]) gab ich eine hinreichende Bedingung für die Lösbarkeit des Systems (64) in der Klasse C unter der alleinigen

Bedingung $\sum_{\nu=1}^{\infty}(\varepsilon_\nu'^2+\varepsilon_\nu''^2)<\infty$. Wie mein Beweis zeigt, folgt aus der Lösbarkeit des Systems (64) in der Klasse K der beschränkten und L-integrierbaren Funktionen mit vorgeschriebenes ε_k', ε_k'' und $\sum_{\nu=1}^{\infty}(\varepsilon_\nu'^2+\varepsilon_\nu''^2)<\infty$ sofort die Lösbarkeit der analogen Problem in der Klasse C. Und zwar gilt dies auch für beliebige Orthogonalsysteme. Im folgenden zeigen wir einen weiteren Reduktionssatz.

Satz VIII. *Wenn das System* (63) *mit* $\sum_{\nu=1}^{\infty}\varepsilon_\nu^2<\infty$ *eine Lösung aus der Klasse* L_p *mit* $p>2$ *besitzt, so besitzt es auch eine Lösung aus der Klasse* C.

Beweis. Wie eine geometrische Betrachtung zeigt (siehe ZYGMUND, *Trigonometrical Series* (Warszawa, 1935), besonders p. 218—219), ist die notwendige und hinreichende Bedingung der Lösbarkeit von (64) in der Klasse K (also auch in der Klasse C), wenn $\sum_{\nu=1}^{\infty}(\varepsilon_\nu'^2+\varepsilon_\nu''^2)<\infty$ gilt, daß

(74a) $$\max_{\alpha_1,\beta_1,\ldots,\alpha_k,\beta_k}|a_1\alpha_1+b_1\beta_1+\ldots+a_k\alpha_k+b_k\beta_k|> \\ >c_1(a_1^2+b_1^2+\ldots+a_k^2+b_k^2)^{1/2},$$

wo $a_1, a_2, \ldots, a_k$; $b_1, b_2, \ldots, b_k$ beliebige reelle Konstanten und $\alpha_1, \beta_1, \ldots, \alpha_k, \beta_k$ die Fourierkonstanten

$$\frac{1}{\pi}\int_0^{2\pi} f(x)\cos n_\nu x\,dx,\quad \frac{1}{\pi}\int_0^{2\pi} f(x)\sin n_\nu x\,dx\qquad (\nu=1,2,\ldots,k)$$

derjenigen Funktionen durchlaufen, für welche $|f(x)|\leq 1$ überall gilt und c_1 eine von den a_ν, b_ν und k unabhängige Konstante bedeutet. Wenn wir das trigonometrische System durch ein allgemeines Orthogonalsystem ersetzen und die Lösbarkeit von (63) in Klasse L_p untersuchen, so folgt ganz ähnlich als notwendige und hinreichende Bedingung:

(74b) $$\max_{\gamma_1,\ldots,\gamma_k}|d_1\gamma_1+\ldots+d_k\gamma_k|>c_2(d_1^2+d_2^2+\ldots+d_k^2)^{1/2},$$

wo $d_1, d_2, \ldots, d_k$ beliebige reelle Konstante, c_2 eine von den d_ν und k unabhängige Konstante bedeuten und die $\gamma_\nu=\int_0^{2\pi} f(x)\varphi_{n_\nu}(x)\,dx$ die verallgemeinerte Fourier-Konstanten einer $f(x)$ bedeuten, für welche $\int_0^{2\pi} |f(x)|^p dx \leq 1$ gilt. Es ist aber

$$\max_{\gamma_1, \ldots, \gamma_k} |d_1\gamma_1 + d_2\gamma_2 + \ldots + d_k\gamma_k| =$$

$$= \operatorname{Max}_f \left| \int_0^{2\pi} f(x) \left[d_1\varphi_{n_1}(x) + d_2\varphi_{n_2}(x) + \ldots + d_k\varphi_{n_k}(x)\right] \right| dx =$$

$$= \left[\int_0^{2\pi} |d_1\varphi_{n_1}(x) + d_2\varphi_{n_2}(x) + \ldots + d_k\varphi_{n_k}(x)|^{\frac{p}{p-1}} dx\right]^{\frac{p-1}{p}}$$

nach der Hölderschen Ungleichung; also ist (74b) mit der Ungleichung

$$\text{(74c)} \quad \int_0^{2\pi} |d_1\varphi_{n_1}(x) + d_2\varphi_{n_2}(x) + \ldots + d_k\varphi_{n_k}(x)|^{\frac{p}{p-1}} dx >$$

$$> c_3 (d_1^2 + d_2^2 + \ldots + d_k^2)^{\frac{p}{2p-2}}$$

($d_1, d_2, \ldots, d_k$ beliebig) äquivalent. Wenn aber $p > 2$ ist, so ist $\frac{p}{p-1} < 2$, also folgt, da allgemein

$$\int_0^{2\pi} |F(x)|^{\frac{p}{p-1}} dx = \int_0^{2\pi} |F(x)|^{\frac{p-2}{p-1}} |F(x)|^{\frac{2}{p-1}} dx <$$

$$< \left[\int_0^{2\pi} |F(x)|\, dx\right]^{\frac{p-2}{p-1}} \left[\int_0^{2\pi} |F(x)|^2 dx\right]^{\frac{1}{p-1}}$$

gilt, aus (74c)

$$c_3(d_1^2 + \ldots + d_k^2)^{\frac{p}{2p-2}} <$$

$$< \left[\int_0^{2\pi} |d_1\varphi_{n_1}(x) + \ldots + d_k\varphi_{n_k}(x)|\, dx\right]^{\frac{p-2}{p-1}} (d_1^2 + \ldots + d_k^2)^{\frac{1}{p-1}}.$$

Dann ist aber

$$\int_0^{2\pi} |d_1\varphi_{n_1}(x) + \ldots + d_k\varphi_{n_k}(x)|\, dx \geqq c_4 (d_1^2 + d_2^2 + \ldots + d_k^2)^{\frac{1}{2}},$$

was eben die notwendige und hinreichende Bedingung dafür ist, daß (63) mit $\sum_{\nu=1}^{\infty} \varepsilon_\nu^2 < \infty$ in der Klasse K (also auch in der Klasse C) lösbar sei.

Es ist wahrscheinlich, daß das System (64) mit $\sum_{\nu=1}^{\infty} (\varepsilon_\nu'^2 + \varepsilon_\nu''^2) < \infty$ in der Klasse C lösbar ist, wenn die Indexfolge $n_1, n_2, \ldots$ die Bedingung $n_{i_1} + n_{i_2} \neq n_{i_3}$ erfüllt. Satz IX wird zeigen, daß der analoge Satz richtig ist, wenn die Klasse C durch die Klasse P der positiven und L-integrablen Funktionen ersetzt wird. Es existieren bekanntlich notwendige und hinreichende Bedingungen dafür, daß eine trigonometrische Reihe die Fourierreihe einer positiven Funktion sei; es scheint mir aber sehr schwer, Satz IX aus diesen abzuleiten. Ein Beispiel für eine die obige Bedingung $n_{i_1} + n_{i_2} \neq n_{i_3}$ erfüllende Folge bildet offenbar die Folge $c, d+c, 2d+c, \ldots, kd+c, \ldots$ mit $c < d$.

Satz IX. *Es sei* $\sum_{\nu=1}^{\infty} (\varepsilon_\nu'^2 + \varepsilon_\nu''^2) < \infty$ *und* $0 < n_1 < n_2 < \ldots$ *eine Indexfolge, für welche immer* $n_{i_1} + n_{i_2} \neq n_{i_3}$ *gilt. Dann ist das System* (64) *in der oben definierten Klasse P lösbar.*

Vor dem Beweis des Satzes IX beweisen wir einen Hilfssatz.

Hilfssatz B. *Wenn die Indexfolge* $0 < n_1 < n_2 < \ldots$ *die obige Bedingung* $n_{i_1} + n_{i_2} \neq n_{i_3}$ *erfüllt und für*

$$f(x) = \sum_{\nu=1}^{K} (a_k \cos n_k x + b_k \sin n_k x)$$

überall

$$-m \leqq f(x) \leqq M, \quad m > 0, \quad M > 0 \tag{75}$$

gültig ist, so ist

$$\left(\sum_{\nu=1}^{K} (a_k^2 + b_k^2)\right)^{1/2} \leqq \sqrt{2} \min (M, m).$$

Beweis des Hilfssatzes B. Es sei ohne Beschränkung der Allgemeinheit $m > M$. Da die Funktionen $M - f(x)$, $f(x) + m$ nichtnegativ sind, gilt nach der Cauchyschen Ungleichung

$$I_1^2 = \left[\int_0^{2\pi} (M-f)^2 (m+f)\, dx\right]^2 =$$

$$= \left[\int_0^{2\pi} (M-f)^{3/2} \{(M-f)^{1/2} (m+f)\}\, dx\right]^2 \leqq \tag{76}$$

$$\leqq \int_0^{2\pi} (M-f)^3\, dx \int_0^{2\pi} (M-f)(m+f)^2\, dx = I_2 I_3.$$

Nun ist

$$I_1^2=\int_0^{2\pi}(M^2-2Mf+f^2)(m+f)\,dx=$$

$$=2\pi M^2m+(m-2M)\pi\sum_{\nu=1}^{K}(a_\nu^2+b_\nu^2)$$

da $\int_0^{2\pi} f(x)\,dx=0$ wegen der des Verschwindens des konstanten Gliedes von $f(x)$ und $\int_0^{2\pi} f^3(x)\,dx=0$ wegen der Indexbedingung $n_{i_1}+n_{i_2}\neq n_{i_3}$. Es folgt ganz ähnlich

$$I_2=2\pi M^3+3\pi M\sum_{\nu=1}^{K}(a_\nu^2+b_\nu^2)$$

und

$$I_3=2\pi m^2M+(M-2m)\pi\sum_{\nu=1}^{K}(a_\nu^2+b_\nu^2).$$

Aus (76) folgt also, wenn wir $\sum_{\nu=1}^{K}(a_\nu^2+b_\nu^2)$ mit S^2 bezeichnen,

$$(2M^2m+(m-2M)S^2)^2\leqq(2M^3+3MS^2)(2m^2M+(M-2m)S^2),$$
$$4M^4m^3+4M^2m(m-2M)S^2+(m-2M)^2S^4\leqq$$
$$\leqq 4M^4m^2+(6M^2m^2+2M^4-4mM^3)S^2+3M(M-2m)S^4,$$

also auch offenbar

$$4M^2m(m-2M)+(m-2M)^2S^2\leqq$$
$$\leqq 6M^2m^2+2M^3(M-2m)+3M(M-2m)S^2,$$
$$(m^2+2Mm+M^2)S^2\leqq 2M^2(m^2+2Mm+M^2),$$
$$S\leqq M\sqrt{2}.$$

B e w e i s d e s S a t z e s IX. Es sei $\eta_1, \eta_2, \ldots$ eine konvexe Nullfolge von der Beschaffenheit, daß die Reihe

$$\sum_{\nu=1}^{\infty}\frac{\varepsilon_\nu'^2+\varepsilon_\nu''^2}{\eta_{n_\nu}^2} \tag{77}$$

noch konvergiert. Wenn wir Hilfssatz B anwenden, so folgt nach einem wohlbekannten Satze von F. RIESZ[10]) die Existenz einer

[10]) F. RIESZ, Sur certain systemes d'équations integrales, *Annales de l'École Normale Superieure*, **28** (1911), S. 33–62.

nichtabnehmenden Funktion $\alpha(x)$, für welche

$$(78)\qquad \frac{1}{\pi}\int_0^{2\pi}\cos n_\nu x\, d\alpha(x)=\frac{\varepsilon'_\nu}{\eta_{n_\nu}},\quad \frac{1}{\pi}\int_0^{2\pi}\sin n_\nu x\, d\alpha(x)=\frac{\varepsilon''_\nu}{\eta_{n_\nu}}.$$

Bekanntlich[11] ist die Reihe $\sum_{\nu=1}^{\infty}\eta_\nu\cos\nu x$ die Fourierreihe einer positiven und L-integrablen Funktion $\psi(x)$; wenn $\sigma_n^+(x)$ das n-te Cesàro-Mittel der Funktion $\psi(x)$, $\sigma_n(x)$ das n-te Cesàro-Mittel der komponierten Reihe von $d\alpha(t)$ und $\psi(x)$ bedeutet, so ist

$$(79)\qquad \sigma_n(x)=\frac{1}{\pi}\int_0^{2\pi}\sigma_n^+(x+t)\,d\alpha(t),$$

$$\int_0^{2\pi}|\sigma_n(x)-\sigma_m(x)|\,dx\leqq\frac{1}{\pi}\int_0^{2\pi}|d\alpha(t)|\int_0^{2\pi}|\sigma_n^+(x+t)-\sigma_m^+(x+t)|\,dx\leqq$$

$$\leqq\frac{1}{\pi}\int_0^{2\pi}|d\alpha(t)|\left[\int_0^{2\pi}|\sigma_n^+(x)|\,dx+\int_0^{2\pi}|\sigma_m^+(x)|\,dx\right]=\frac{2}{\pi}\eta_0(\alpha(2\pi)-\alpha(0)),$$

also nach dem bekannten Satze[12] ist die komponierte Reihe die Fourierreihe einer L-integrablen Funktion, welche nach (79) positiv ist, qu. e. d.

Bisher haben wir die Lösbarkeit des Systems (64) in einigen Funktionenklassen untersucht, wenn nur $\sum_{\nu=1}^{\infty}(\varepsilon_\nu'^2+\varepsilon_\nu''^2)<\infty$; von der Lösung selbst wußten wir nichts anderes, als daß sie zu der geforderten Klasse gehört. Wenn speziell die Indexfolge durch $n_k=4^{2^k}$ $(k=1,2,\ldots)$ gegeben ist, dann können wir — wieder $\sum_{\nu=1}^{\infty}(\varepsilon_\nu'^2+\varepsilon_\nu''^2)<\infty$ vorausgesetzt — die Existenz einer Lösung von (64) im Klasse C sichern, welche nur „wenige" nichtverschwindende Fourierkoeffizienten besitzt. Genauer:

Satz X. *Es sei* $n_k=4^{2^k}$ $(k=1,2,\ldots)$ *und* $\sum_{\nu=1}^{\infty}(\varepsilon_\nu'^2+\varepsilon_\nu''^2)$ *vorgegeben. Dann hat des System* (64) *eine solche Lösung, bei*

---

[11]) Siehe A. Zygmund, *Trigonometrical Series* (Warszawa—Lwów, 1935), S. 109.

[12]) Satz von Steinhaus-Gross. Siehe Zygmund, *Trigonometrical Series*, S. 84.

welcher für die Indizes $N_1 < N_2 < \ldots < N_k < \ldots$ der nichtverschwindenden Fourierkoeffizienten

$$N_k^{1/k} \geqq q > 1$$

gilt, unabhängig von k.

Beweis des Satzes X. Nach dem Satze der Arbeit [8]) ist das System

$$(80) \qquad \frac{1}{\pi}\int_0^{2\pi} f(x)\cos 4^{2^i} x\,dx = \varepsilon_i', \quad \frac{1}{\pi}\int_0^{2\pi} f(x)\sin 4^{2^i} x\,dx = \varepsilon_i',$$

mit $\sum_{\nu=1}^{\infty}(\varepsilon_\nu'^2 + \varepsilon_\nu''^2) < \infty$ in der Klasse C lösbar; diese Lösung sei $f_1(x) \sim \sum_{\nu=1}^{\infty}(a_\nu \cos \nu x + b_\nu \sin \nu x)$. Wir betrachten diejenige trigonometrische Reihe $\sum_{\nu=1}^{\infty} \sigma_\nu \cos \nu y$, welche durch formale Ausmultiplizieren des Produktes $\prod_{i=1}^{\infty}(1 + \cos 4^{2^i} y)$ entsteht. Die Partialprodukte $\prod_{i=1}^{K}(1 + \cos 4^{2^i} y)$ sind offenbar nichtnegativ und wegen der Seltenheit der Indizes haben sie nach Multiplikation und Umordnung die Form $\sum_{\nu=0}^{M} c_\nu \cos \nu y$, wo $M = 4^{2^1} + 4^{2^2} + \ldots + 4^{2^K}$. Dann ist aber nach einem bekannten Satze[13]) die Reihe $\sum_{\nu=1}^{\infty} c_\nu \cos \nu y$ eine Fourier—Stieltjes-Reihe und nach einem ebenfalls bekannten Satze[14]) die komponierte Reihe $\sum_{\nu=0}^{\infty} c_\nu (a_\nu \cos \nu x + b_\nu \sin \nu x)$ die Fourierreihe einer überall stetigen Funktion $f_2(x)$. Diese Funktion erfüllt offenbar (80), die nichtverschwindenden Fourierkoeffizienten haben die Indizes $4^{2^{i_1}} + 4^{2^{i_2}} + \ldots + 4^{2^{i_r}}$, also, wenn $N_1 < N_2 < \ldots$ die der Größe nach geordnete Folge dieser Indizes bedeutet, so ist offenbar $N_k > 2^k$ bei jeder k erfüllt.

Wie ich früher bewiesen hatte, ist das System (64) in der Klasse L der nach Lebesgue integrierbaren Funktionen lösbar, falls nur die vorgegebenen ε_ν', ε_ν'' eine Nullfolge bilden und die Indizes

[13]) ZYGMUND, *Trigonometrical Series*, S. 87—88.

[14]) ZYGMUND, *Trigonometrical Series*, S. 100.

n_k die Bedingung $\frac{n_{k+1}}{n_k} \geqq q > 1$ erfüllen. Es ist naheliegend zu fragen, ob man nicht ähnlicherweise die Lösbarkeit in der Klasse L_p, $1 < p < 2$ sichern kann. Eine Möglichkeit wäre für den Zahlen ε_ν, ε''_ν eine Bedingung von der Art $\sum_{\nu=1}^{\infty} (|\varepsilon'_\nu|^\alpha + |\varepsilon''_\nu|^\alpha)$ zu fordern, wo α nur von p abhängt; das kommt aber nach einem Satze von A. ZYGMUND[15]) nicht im Betracht, wie auch die Indexfolge $n_1, n_2, \ldots$ sei. Man könnte aber denken, daß, wenn wir die Zahlen ε'_ν und ε''_ν der Voraussetzung

$$(81) \qquad |\varepsilon'_\nu| \leqq \frac{A}{\nu^\delta}, \ |\varepsilon''_\nu| \leqq \frac{A}{\nu^\delta} \qquad (\delta < 1/2, \ \delta = \delta(p))$$

unterwerfen und die Indizes $n_1, n_2, \ldots$ „genügend zerstreut" wählen, dann gibt es stets eine $f(x)$ in L_p $(1 < p < 2)$ mit $f(x) \sim \sum_{\nu=0}^{\infty} (a_\nu \cos \nu x + b_\nu \sin \nu x)$ so, daß $a_{n_k} = \varepsilon'_k$, $b_{n_k} = \varepsilon''_k$ $(k = 1, 2, \ldots)$ gilt. Wir zeigen, daß auch dies nicht der Fall ist. Wenn nämlich solche p mit $1 < p < 2$, δ und Indexfolge $n_1 < n_2 < \ldots$ existieren, so bilden wir die Reihe

$$(82) \qquad \sum_{k=2}^{\infty} \frac{1}{\sqrt{k} \log k} \cos n_k x \sim g_1(x).$$

Da die Quadratsumme der Koeffizienten in (82) konvergiert, existiert nach einem Satze von LITTLEWOOD[16]) eine Folge $\eta_1, \eta_2, \ldots$ mit $|\eta_k| = 1$ $(k = 1, 2, \ldots)$ so, daß die Reihe $\sum_{k=2}^{\infty} \frac{\eta_k}{\sqrt{k} \log k} \cos n_k x$ die Fourierreihe einer Funktion $g_2(x)$ aus der Klasse $L_{\frac{p}{p-1}}$ ist. Nach einem Satze von KACZMARZ[17]) ist diejenige trigonometrische Reihe, welche durch Komposition der Fourierreihe zweier Funktionen aus der Klasse L_p bzw. $L_{\frac{p}{p-1}}$ entsteht, die Fourierreihe einer beschränkten Funktion $g_3(x)$. Es sei nun $\varepsilon'_\nu = \frac{\overline{\eta}_\nu}{\nu^\delta}$, $\varepsilon''_\nu = \frac{\overline{\eta}_\nu}{\nu^\delta}$. Dann ist (81) erfüllt und wenn unsere Behauptung richtig wäre, exi-

[15]) A ZYGMUND, On the convergence of lacunary trigonometrical series, *Fundamenta Math.*, 16 (1930), S. 90–107.

[16]) J. E. LITTLEWOOD, On the mean-values of power series, *Proceedings London Math. Society*, 25 (1926), S. 328–337.

[17]) S. KACZMARZ, On some classes of Fourier-series, *Journal London Math. Society*, 8 (1933), S. 39–46.

stierte eine Funktion $g_4(x)$ in L_p mit

$$g_4(x) \sim \sum_{\nu=0}^{\infty} (a_\nu \cos \nu x + b_\nu \sin \nu x),\; a_{n_\nu} = \frac{\bar{\eta}_\nu}{\nu^\delta},\; b_{n_\nu} = \frac{\bar{\eta}_\nu}{\nu^\delta}.$$

Wenn wir den obenerwähnten Satz von KACZMARZ auf die Funktionen $g_2(x)$ und $g_4(x)$ anwenden, gelangen wir zu dem Ergebnis, daß die Reihe $\sum_{k=2}^{\infty} \frac{1}{k^{1/2+\delta}\log k} \cos n_k x$ die Fourierreihe einer beschränkten Funktion ist, also auch die arithmetischen Mittel erster Ordnung dieser Reihe überall beschränkt sind. Das ist aber für $x=0$ augenscheinlich nicht der Fall.

§. 3.

Der wohlbekannte Satz von YOUNG—HAUSDORFF wirft Licht auf den Zusammenhang des Konvergenzexponenten der Fourierkoeffizienten und des Integrierbarkeitsexponenten der dargestellten Funktion $f(x)$; gehört z. B. $f(x)$ der Klasse L_p, $1 < p < 2$, so ist die Reihe $\sum_{\nu=1}^{\infty} (|a_\nu|^{\frac{p}{p-1}} + |b_\nu|^{\frac{p}{p-1}})$ konvergent und der Exponent $\frac{p}{p-1}$ läßt sich im allgemeinen nicht verkleinern. ZYGMUND[15]) und PALEY[18]) bemerkten, daß, wenn man nicht alle Koeffizienten betrachtet, sondern nur gewisse a_{n_k} und b_{n_k}, wobei $\frac{n_{k+1}}{n_k} \geqq q > 1$, so kann man viel mehr aussagen, nämlich daß $\sum_{\nu=1}^{\infty} (a_{n_\nu}^2 + b_{n_\nu}^2)$ konvergiert. Wie dieselben Verfasser später zeigten, konvergiert dann merkwürdigerweise auch die Reihe $\sum_{\nu=1}^{\infty} (|a_{n_\nu}| + |b_{n_\nu}|)$ und zwar auch in dem Falle, wenn die Funktion $f(x)$ und ihre Konjugierte zu der Klasse L gehören.

Die Indexfolge n_i ist sehr selten; wir betrachten anstatt deren solche Indexfolgen $n_1 < n_2 < \ldots$, wie wir sagen, B_l-Folgen, für welchen die Lösungzahl der Gleichung

(83) $$n_{i_1} + n_{i_2} + \ldots + n_{i_l} = N$$

unter einer von N unabhängigen Schranke bleibt. Diejenigen In-

---

[18]) R. E. A. C. PALEY, On the lacunary coefficients of power series, *Annals of Math.*, 34 (1933), S. 615–616.

dexfolgen $j_1 < j_2 < \ldots$, für welche $\frac{j_{i+1}}{j_i} \geqq q > 1$ gilt, gehören auch zu den B_l-Folgen bei jeder ganzen $l \geqq 2$. Wie wir schon bemerkten, können solche B_l-Folgen viel dichter sein, als eine solche Folge $j_1, j_2, \ldots$ Betrachten wir einfachheitshalber den Fall $l = 2$ und untersuchen die dichteste Folge $1 = d_1 < d_2 < \ldots$ der ganzen Zahlen, für welche die Lösungzahl $d_i + d_k = N$ höchstens 1 ist. Offenbar ist die Anzahl der verschiedenen der Summen $d_i + d_k$ $(1 \leqq i \leqq k \leqq n-1)$ höchstens $\frac{n(n-1)}{2}$, also gilt

$$d_n \leqq d_{n-1} + \frac{n(n-1)}{2};$$

hieraus folgt durch Addition $d_n \leqq \frac{n^3}{6} + 1$. Wahrscheinlich existieren B_2-Folgen $d_1' < d_2' < \ldots$ mit $d_n' = O(n^{2+\varepsilon})$. Es sei immer $l \geqq 2$ und ganz. Wir beweisen die folgenden Sätze.

Satz XI. *Es sei die Indexfolge* $n_1 < n_2 < \ldots$ *eine* B_l*-Folge,* $q > 2$, $f(x) \in L_{\frac{lq}{lq-1}}$ *und* $f(x) \sim \sum\limits_{\nu=0}^{\infty} (a_\nu \cos \nu x + b_\nu \sin \nu x)$. *Dann ist* $\sum\limits_{k=1}^{\infty} |a_{n_k}|^q < \infty$.

Satz XII. *Es sei* $n_1 < n_2 < \ldots$ *eine* B_l*-Folge,* $q > 2l$ *und* $g(x) \in L_{\frac{q}{q-1}}$, *wo* $g(x) \sim \sum\limits_{\nu=0}^{\infty} (a_\nu \cos \nu x + b_\nu \sin \nu x)$. *Dann ist* $\sum\limits_{k=1}^{K} (a_{n_k}^2 + b_{n_k}^2) = O\left(K^{1-\frac{2l}{q}}\right)$.

Bevor wir diese Sätze beweisen, schicken wir zwei Hilfssätze voraus.

Hilfssatz C. *Wenn* $\sum\limits_{k=1}^{\infty} |\beta_k|^{\frac{q}{q-1}} < \infty$, $q \geqq 2$ *und* $n_1, n_2, \ldots$ *eine* B_l*-Folge bedeutet, dann gehört die Funktion* $g(x) \sim \sum\limits_{k=1}^{\infty} \beta_k e^{i n_k x}$ *zu der Klasse* H_{lq} *und* $\int\limits_0^{2\pi} |g(x)|^{lq} dx \leqq c_6(l, q) \left(\sum\limits_{k=1}^{\infty} |\beta_k|^{\frac{q}{q-1}}\right)^{l(q-1)}$

Vorbemerkung. Einen in ähnlicher Richtung wie Hilfssatz C liegenden Satz für die spezielleren Reihe $\sum_{k=1}^{\infty}(a_k \cos m_k x + b_k \sin m_k x)$ mit $\frac{m_{k+1}}{m_k} \geq q > 1$ findet man in einer Note des Herrn A. ZYGMUND[19]).

Beweis des Hilfssatzes C. Es sei $\left(\sum_{k=1}^{\infty} \beta_k z^{n_k}\right)^l = \sum_{m=1}^{\infty} A_m z^m$. Wenn wir beweisen, daß $\sum_{m=1}^{\infty} |A_m|^{\frac{q}{q-1}}$ konvergiert, dann ist nach dem Satze von YOUNG—HAUSDORFF $[g(x)]^l \in H_q$. Es gilt offenbar

$$(84) \qquad A_m = \sum_{n_{i_1}+\ldots+n_{i_l}=m} \beta_{i_1}\beta_{i_2}\ldots\beta_{i_l}.$$

Dann ist aber

$$|A_m|^{\frac{q}{q-1}} = \left|\sum_{n_{i_1}+\ldots+n_{i_l}=m} \beta_{i_1}\beta_{i_2}\ldots\beta_{i_l}\right|^{\frac{q}{q-1}} \leq$$

$$\leq c_7(l, q) \sum_{n_{i_1}+\ldots+n_{i_l}=m} |\beta_{i_1}|^{\frac{q}{q-1}} \ldots |\beta_{i_l}|^{\frac{q}{q-1}},$$

da die erste Summe nach der Voraussetzung nur Glieder enthält, dessen Anzahl unter einer von m unabhängigen Schranke bleibt. Dann ist aber

$$\sum_{m=1}^{\infty} |A_m|^{\frac{q}{q-1}} \leq c_7(l, q)\left(\sum_{\nu=1}^{\infty} |\beta_\nu|^{\frac{q}{q-1}}\right)^l, \text{ qu. e. d.}$$

Hilfssatz D. *Wenn* $q \geq 2l$, $z = e^{i\varphi}$, $n_1, n_2, \ldots$ *die obige Indexfolge und* $P(z) = \sum_{k=1}^{K} \gamma_k z^{n_k}$, *dann ist*

$$\int_{|z|=1} |P(z)|^q \, d\varphi \leq c_8(l, q) K^{\frac{q}{2}-l} \left(\sum_{k=1}^{K} |\gamma_k|^2\right)^{\frac{q}{2}}.$$

Beweis. Es sei $q' < q < q''$, wo q' und q'' gerade ganze Zahlen bedeuten, und es sei vorausgesetzt, daß Hilfssatz D für gerade Exponenten schon bewiesen sei. Wenn wir die Zahlen δ_1 und δ_2 durch

$$\delta_1 + \delta_2 = 1, \; q'\delta_1 + q''\delta_2 = q$$

definieren, so folgt aus der Hölderschen Ungleichung

[19]) A. ZYGMUND, Note on trigonometrical and Rademacher series, *Prace mat. fiz.*, 44 (1936), S. 91—107.

$$\int\limits_{|z|=1} |P(z)|^q d\varphi = \int\limits_{|z|=1} |P(z)|^{\delta_1 q'} |P(z)|^{\delta_2 q''} d\varphi <$$
$$< \Big(\int\limits_{|z|=1} |P(z)|^{q'} d\varphi\Big)^{\delta_1} \Big(\int\limits_{|z|=1} |P(z)|^{q''} d\varphi\Big)^{\delta_2} <$$
$$< c_8(l,q)^2 K^{\left(\frac{q'}{2}-l\right)\delta_1 + \left(\frac{q''}{2}-l\right)\delta_2} \Big(\sum_{k=1}^{K} |\gamma_k|^2\Big)^{\frac{1}{2}(q'\delta_1 + q''\delta_2)} =$$
$$= c_9(l,q) K^{\frac{q}{2}-l} \Big(\sum_{k=1}^{K} |\gamma_k|^2\Big)^{\frac{q}{2}},$$

wie behauptet. Es genügt also Hilfssatz D für gerade q zu beweisen. Aus der Definition der Indexfolge folgt offenbar, daß die Lösungzahl der Gleichung $n_{i_1} + n_{i_2} + \ldots + n_{i_{q/2}} = N$ kleiner ist als $c_{10}(l,q) N^{\frac{q}{2}-l}$. Aus dieser Bemerkung und aus der Cauchyschen Ungleichung folgt, wenn wir $P(z)^{\frac{q}{2}} = \sum_{n=1}^{K^{\frac{q}{2}}} A_n z^n$ setzen,

$$\frac{1}{2\pi} \int\limits_{|z|=1} |P(z)|^q d\varphi = \sum_{n=1}^{K^{\frac{q}{2}}} |A_n|^2 = \sum_{n=1}^{K^{\frac{q}{2}}} \Big| \sum_{n_{i_1}+\ldots+n_{i_{q/2}}=n} \gamma_{i_1}\gamma_{i_2}\ldots\gamma_{i_{q/2}} \Big|^2 \leqq$$
$$\leqq c_{11}(l,q) \sum_{n=1}^{K^{\frac{q}{2}}} n^{\frac{q}{2}-l} \sum_{n_{i_1}+\ldots+n_{i_{q/2}}=n} |\gamma_{i_1}|^2 |\gamma_{i_2}|^2 \ldots |\gamma_{i_{q/2}}|^2 \leqq$$
$$\leqq c_{12}(l,q) K^{\frac{q}{2}-l} \Big(\sum_{\nu=1}^{K} |\gamma_\nu|^2\Big)^{\frac{q}{2}}.$$

Beweis des Satzes XI. Es sei

$$f(x) \sim \sum_{\nu=0}^{\infty} (a_\nu \cos \nu x + b_\nu \sin \nu x),$$

$$\Phi_K(z) = \sum_{k=1}^{K} (a_{n_k}^{q-1} \operatorname{sign} a_{n_k}^q \cos n_k x + b_{n_k}^{q-1} \operatorname{sign} b_{n_k}^q \sin n_k x).$$

Dann ist

$$S = \sum_{k=1}^{K} (|a_{n_k}|^q + |b_{n_k}|^q) = \frac{1}{\pi} \int_0^{2\pi} f(x)\, \Phi_k(x)\, dx \leqq$$
$$\leqq \frac{1}{\pi} \Big[\int_0^{2\pi} |f(x)|^{\frac{lq}{lq-1}} dx\Big]^{\frac{lq-1}{lq}} \Big[\int_0^{2\pi} |\Phi_k(x)|^{lq} dx\Big]^{\frac{1}{lq}}.$$

Wenn man Hilfssatz C mit $\Re g(x) = \Phi_K(x)$ anwendet, folgt

$$S \leqq \frac{1}{\pi} \left[\int_0^{2\pi} |f(x)|^{\frac{lq}{lq-1}} dx \right]^{\frac{lq-1}{lq}} c_6(l, q) \left(\sum_{k=1}^{\infty} |a_{n_k}^{q-1} \operatorname{sign} a_{n_k}^q + \right.$$

$$\left. + b_{n_k}^{q-1} \operatorname{sign} b_{n_k}^q |^{\frac{q}{q-1}} \right)^{\frac{q-1}{q}} \leqq c_{13}(l, q) \left[\int_0^{2\pi} |f(x)|^{\frac{lq}{lq-1}} dx \right]^{\frac{lq-1}{lq}} S^{\frac{q-1}{q}},$$

also

$$S < c_{14}(l, q) \left[\int_0^{2\pi} |f(x)|^{\frac{lq}{lq-1}} dx \right]^{\frac{lq-1}{l}}$$

unabhängig von K, qu. e. d.

Beweis des Satzes XII. Es sei

$$U = \sum_{k=1}^{K} (a_{n_k}^2 + b_{n_k}^2) = \frac{1}{\pi} \int_0^{2\pi} g(x) \left[\sum_{k=1}^{K} (a_{n_k} \cos n_k x + b_{n_k} \sin n_k x) \right] dx,$$

also nach der Hölderschen Ungleichung

$$U \leqq \frac{1}{\pi} \left[\int_0^{2\pi} |g(x)|^{\frac{q}{q-1}} dx \right]^{\frac{q-1}{q}} \left[\int_0^{2\pi} \left| \sum_{k=1}^{K} (a_{n_k} \cos n_k x + b_{n_k} \sin n_k x) \right|^q dx \right]^{\frac{1}{q}}.$$

Wenn wir Hilfssatz D mit $\Re P(z) = \sum_{k=1}^{K} (a_{n_k} \cos n_k x + b_{n_k} \sin n_k x)$ anwenden, folgt, daß

$$U < c_{14}(l, q) \left(\int_0^{2\pi} |g(x)|^{\frac{q}{q-1}} dx \right)^{\frac{q-1}{q}} K^{\frac{1}{2} - \frac{l}{q}} U^{\frac{1}{2}},$$

also

$$U < c_{15}(l, q) K^{1 - \frac{2l}{q}} \left(\int_9^{2\pi} |g(x)|^{\frac{q}{q-1}} dx \right)^{\frac{2q-2}{q}}.$$

Satz XIII entspricht dem oben erwähnten Satz von ZYGMUND. In diesem Satze figuriert die Voraussetzung, daß $f(x)$ und $\bar{f}(x)$ zu der Klasse L gehören. Im Satz XIII fordern wir nur die L-Integrierbarkeit von $f(x)$ selbst, und statt der anderen Hälfte der Zygmundschen Voraussetzung tritt die Forderung, daß die Reihe „lückenartig" sei. Genauer:

Satz XIII. *Es sei* $f(x) \in L$, $f(x) \sim \sum_{\nu=0}^{\infty} (a_\nu \cos \nu x + b_\nu \sin \nu x)$, *die Indexfolge* $n_1 < n_2 < \dots$ *sei von der Beschaffenheit, daß* $\frac{n_{k+1}}{n_k} \geqq q > 1$ *und es gäbe ein* q' *mit* $q > q'^2 > 1$ *so, daß* $a_n = b_n = 0$ *für alle* n, *für welche*

$$\frac{n_k}{q'} \leqq n \leqq n_k q', \; n \neq n_k \qquad (k = 1, 2, \dots).$$

Dann ist die Reihe $\sum_{k=1}^{\infty} (a_{n_k}^2 + b_{n_k}^2)$ *konvergent.*

Beweis. Es sei l die kleinste ganze Zahl, für welche

$$1 + \frac{1}{q^{l-1} - 1} < q'$$

gilt. Es sei $\varepsilon_1, \varepsilon_2, \dots$ eine beliebige Folge aus Zahlen ± 1, $0 \leqq l < r$ und wir betrachten die formal gebildete trigonometrische Reihe

$$\sum_{\nu=1}^{\infty} c_\nu^{(l,r)} \cos \nu x \sim \prod_{s=1}^{\infty} (1 + \varepsilon_{ls+r} \cos n_{ls+r} x).$$

Ein analoger Gedankengang, wie beim Satz X, ergibt, daß alle diese Reihen Fourier—Stieltjes-Reihen sind. Dann sind aber alle Reihen, welche aus der Fourierreihe der L-integrierbaren Funktion $f(x)$ durch Komposition mit den Reihen $\sum_{\nu=1}^{\infty} c_\nu^{(l,r)} \cos \nu x$ entstehen, nach einem bekannten Satze[20]) wieder Fourierreihen von L-integrierbaren Funktionen. Diese Reihen sind aber wegen der Indexbedingungen von der Form

$$\sum_{s=1}^{\infty} \varepsilon_{ls+r} \, (a_{n_{ls+r}} \cos n_{ls+r} x + b_{n_{ls+r}} \sin n_{ls+r} x).$$

Wie ich aber früher bewies[21]), ist die Quadratsumme der Koeffizienten einer trigonometrischen Reihe konvergent, wenn die Reihe von der Beschaffenheit ist, daß nach Komposition mit einer beliebigen Folge aus Zahlen ± 1 immer eine Fourierreihe einer L-integrierbaren Funktion entsteht; es ist also $\sum_{s=1}^{\infty} (a_{n_{ls+r}}^2 + b_{n_{ls+r}}^2)$ konvergent $(r = 0, 1, \dots, l-1)$, also auch die Reihe $\sum_{s=1}^{\infty} (a_{n_s}^2 + b_{n_s}^2)$.

---

[20]) Zygmund, *Trigonometrical Series*, S. 100.

[21]) S. Sidon, Ein Satz über die Fourierschen Reihen stetiger Funktionen, *Math. Zeitschrift*, 34 (1932), S. 486.

Bezüglich Hilfssatz D bemerken wir folgendes. Es sei eine B_l-Folge $n_1 < n_2 < \dots$ gegeben mit der Zusatzbedingung, daß $\frac{n_{2k}}{n_k} \geqq a > 1$. Wenn $N_1 < N_2 < \dots$ die der Größe nach geordnete Summen von der Gestalt $n_{i_1} + n_{i_2} + \dots + n_{i_j}$ bedeuten (j fest und $1 \leqq j < l$) und $q \geqq 2l$, dann folgt analog

$$(85) \qquad \int_0^{2\pi} \left| \sum_{k=1}^{K} a_k z^{N_k} \right|^q d\varphi < c_{16}(q) K^{\frac{q}{2} - \frac{l}{j}} \left(\sum_{k=1}^{K} |a_k|^2 \right)^{\frac{q}{2}}.$$

Ob diese Ungleichung auch für $2l > q > \frac{2l}{j}$ gilt, konnte ich nicht entscheiden.

Wenn wir den Begriff der B_l-Folgen auf andere Orthogonalsysteme singemäß übertragen wollen, müssen wir ihn etwas abändern. Bezüglich einem Orthogonalsystem $\varphi_0(x), \varphi_1(x), \dots$ nenne ich eine Indexfolge $n_1 < n_2 < \dots$ eine B_l'-Folge, wenn die verallgemeinerten Fourierkoeffizienten der Funktionenfolge $\Phi_\nu(x) = \left[\sum_{\mu=1}^{\nu} \varphi_{n_\mu}(x) \right]^l$ ($\nu = 1, 2, \dots$) (bezüglich dem System $\varphi_0, \varphi_1, \dots$) gleichmäßig beschränkt sind. Für das trigonometrische System und für $l = 2$ sind also die B_2'-Folgen mit den obenerwähnten B_2^+-Folgen identisch. Für das Walsh-System kann man den Beweis des Hilfssatzes D übertragen, also gilt für jede B_l'-Folge $\nu_1 < \nu_2 < \dots$, $f(x) \sim \sum_{k=1}^{K} a_k \varphi_{\nu_k}(x)$ die Ungleichung

$$(86) \qquad \int_{-1}^{+1} [f(x)]^{2l} dx < c_{17} \left(\sum_{k=1}^{K} a_k^2 \right)^l.$$

Ferner gilt, wenn $M_1 < M_2 < \dots$ diejenige ganzen Zahlen bedeuten, welche im dyadischen System aus h Ziffern bestehen, für das Walsh-System die Ungleichung

$$(87) \qquad \int_{-1}^{+1} \left| \sum_{j=1}^{K} a_j \varphi_{M_j}(x) \right|^p dx \leqq c_{18}(p) \left(\sum_{j=1}^{K} a_j^2 \right)^{\frac{p}{2}}.$$

§. 4.

H. Steinhaus warf die folgende interessante Frage auf. Es sei bekannt, daß sämtliche Partialsummen der Reihe $1+\sum_{\nu=1}^{\infty} d_\nu \cos \nu x$ überall nichtnegativ sind. Folgt aus dieser Eigenschaft, daß $\lim\limits_{n\to\infty} a_n = 0$? Diese Frage ist bisher unentschieden. J. Schur und J. von Neumann bewiesen, daß wenigstens $\liminf\limits_{n\to\infty} a_n = 0$. P. Erdős bewies[22]) schärfer, daß die Anzahl der „großen" Koeffizienten „sehr klein" ist, d. h. die Anzahl derjenigen Koeffizienten a_ν, für welche $\nu \leqq n$ und $|a_\nu| \geqq \varepsilon$ gilt, ist beim festen ε kleiner als $c_{19}(\varepsilon)(\log n)^{\frac{4}{\varepsilon^2}}$. Wenn wir statt der Partialsummen arithmetische Mittel erster Ordnung nehmen, so folgt $\lim\limits_{n\to\infty} a_n = 0$ nicht, wie die Fejérsche formale Reihe

$$1+2\cos x+2\cos 2x+\ldots$$

zeigt. Wenn wir die Partialsummen von $1+\sum_{\nu=1}^{\infty} a_\nu \cos \nu x$ mit $s_k(x)$ bezeichnen, so gilt offenbar

$$(88)\qquad \frac{1}{2\pi}\int_0^{2\pi} |s_k(x)|\, dx = \frac{1}{2\pi}\int_0^{2\pi} s_k(x)\, dx = 1 \qquad (k=0,1,2,\ldots),$$

also sind die geometrischen Integrale der Partialsummen beschränkt. Es ist die Frage naheliegend, ob aus dieser schwächeren Voraussetzung schon nicht $\lim\limits_{n\to\infty} a_n = 0$ folgt. In einer Fußnote einer früheren Arbeit[23]) habe ich behauptet, daß die Antwort auf diese schwächere Frage verneinend ist, d. h. es existiert eine Reihe $\sum_{\nu=1}^{\infty} a_\nu \cos \nu x$ mit den Eigenschaften

$$\int_0^{2\pi} |s_k(x)|\, dx \leqq c_{19} \quad (k=0,1,2,\ldots), \quad \overline{\lim_{n\to\infty}}\, a_n > 0.$$

Wie ich später bemerkte, ergibt mein Beweis nicht diese Be-

---

[22]) P. Erdős, On a conjecture of Steinhaus, *Revista de la Universidad Nacional de Tucuman*, 1 (1940), S. 217—220.

[23]) S. Sidon, Bemerkungen über Fourier- und Potenzreihen, *diese Acta*, 7 (1934), S. 89.

hauptung, sondern nur, daß eine *Umordnung* der Funktionenfolge $1, \cos x, \cos 2x, \ldots$ existiert so, daß für die „umgeordneten" Partialsummen $s_k^+(x)$ gilt $\int_0^{2\pi} |s_k^+(x)|\, dx < c_{19}$ $(k = 0, 1, \ldots)$ und $\overline{\lim\limits_{n\to\infty}}\, a_n > 0$. Für das Walsh-System kann ich aber die obige Behauptung ohne Umordnung beweisen.

Satz XIV. *Wenn $\psi_0(x) = 1$, $\psi_1(x), \ldots$ die Funktionen des Walshschen Orthogonalsystems bedeuten, so gibt es eine Reihe $\sum\limits_{\nu=0}^{\infty} a_\nu \psi_\nu(x)$ so, daß*

$$\int_0^1 \left| \sum_{i=0}^{n} a_i \psi_i(x) \right| dx < c_{20}, \quad \overline{\lim}\, a_n > 0.$$

Beweis. Wir betrachten das Produkt (in voller Allgemeinheit benötigen wir es nur später)

$$P_k(t, x, \varrho) = \prod_{i=1}^{k} (1 + \varrho \psi_{2^i}(t) \psi_{2^i}(x)), \tag{89a}$$

wo ϱ eine feste Zahl bedeutet. Nach Multiplikation erhalten wir offenbar

$$P_k(t, x, \varrho) = 1 + \sum_{n=1}^{2^{k+1}-1} \varrho^j \psi_n(t) \psi_n(x), \tag{89b}$$

wo j die Zifferanzahl von n im dyadischen System bedeutet, also

$$n = 2^{k_1} + 2^{k_2} + \ldots + 2^{k_j}. \tag{89c}$$

Es ist offenbar für $|\varrho| \leq 1$

(89d) $P_k(t, x, \varrho) \geq 0$, Koeff. $\psi_{2^k}(t)$ in $P_k(t, x, \varrho) = \varrho \psi_{2^k}(x)$.

Wenn man von $P_k(t, x, \varrho)$ zu $P_{k+1}(t, x, \varrho)$ übergeht, so bleiben die Koeffizienten von $\psi_n(t) \psi_n(x)$, $n \leq 2^{k+1} - 1$ unverändert. Wir bilden formal die Reihe, für welche die $(2^{k+1} - 1)$-te Partialsumme durch

$$\sum_{\nu=0}^{2^{k+1}-1} a_\nu \psi_\nu(x) \psi_\nu(t) = \prod_{\nu=1}^{k} (1 + \varrho \psi_{2^\nu}(x) \psi_{2^\nu}(t)) \quad k = 0, 1, 2, \ldots \tag{89e}$$

gegeben sind, x beliebig aber fest. Dann ist wegen (89d)

$$a_{2^k} = \varrho, \qquad k = 1, 2, \ldots,$$

also wir haben nur die erste Behauptung des Satzes zu beweisen. Es sei $2^k \leq n < 2^{k+1}$ und $s_n(t, x, \varrho)$ die n-te Partialsumme von (89e);

dann ist offenbar

(90a) $$s_n(t,x,\varrho)=P_{k-1}(t,x,\varrho)+\varrho s_{n-2^k}(t,x,\varrho)\,\psi_k(x)\,\psi_k(t),$$

also wegen $P_{k-1}(t,x,\varrho)\geqq 0$ offenbar

(90b) $$\int_0^1 |s_n(t,x,\varrho)|\,dt\leqq 1+|\varrho|\int_0^1 |s_{n-2^k}(t,x,\varrho)|\,dt.$$

Aus (90b) folgt nach Iteration, falls $|\varrho|<1$

(91) $$\int_0^1 |s_n(t,x,\varrho)|\,dt<1+|\varrho|+|\varrho|^2+\ldots=\frac{1}{1-|\varrho|},$$

also ist Satz XIV bewiesen.

Man kann übrigens mit den obigen Polynomen $P_k(t,x,-1)$ auch den folgenden Satz beweisen.

Satz XV. *Wenn $f(x)$ einseitig beschränkt ist und in ihrer Walsh-Entwicklung alle Koeffizienten verschwinden, für welche j (siehe (89c)) gerade ist, dann ist $f(x)$ beiderseits beschränkt.*

Beweis. Es folgt offenbar aus (89b)

$$\int_0^1 P_k(t,x,\varrho)\,\psi_n(t)\,dt=\varrho^j\,\psi_n(x).$$

Wir können ohne Beschränkung der Allgemeinheit voraussetzen, daß $f(x)\leqq M$. Dann ist nach der Voraussetzung

$$-\sum_{n=1}^{2^{k+1}-1} a_n\psi_n(x)=\int_0^1 f(t)\,P_k(t,x,-1)\,dt\leqq M$$

wegen (89d).

RADEMACHER[24]) bewies, daß, wenn in der Walsh-Entwicklung einer überall stetigen Funktion nur die Koeffizienten a_{2^k} $(k=0, 1, 2,\ldots)$ (also die Glieder mit $j=1$) nicht verschwinden, so ist $\sum_{k=1}^{\infty}|a_{2^k}|<\infty$, ähnlich wie bei dem trigonometrischen System. Dies läßt sich nicht für den Fall ausdehnen, wo wir nur wissen, daß alle Koeffizienten mit $j\geqq 3$ verschwinden. Wir beweisen, daß in diesem Falle die Reihe in jedem Intervalle gleichmäßig konvergiert und es gilt sogar der

---

[24]) H. RADEMACHER, Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen, *Math. Annalen*, **87** (1922), S. 112—138.

Satz XVI. *Die Walsh-Entwicklung einer überall stetigen Funktion $f(x)$ habe die Eigenschaft, daß es eine Zahl H gebe so, daß alle solche Koeffizienten a_n verschwinden, deren Index wenigstens $(H+1)$-ziffrig im dyadischen System ist, d. h. $j \geqq H+1$ (siehe (89b)). Dann konvergiert die Reihe in jedem Intervalle gleichmäßig.*

Beweis. Wir betrachten mit den vorigen Bezeichnungen (s. Satz XIV) den Ausdruck

$$(92)\quad \alpha_1 s_n(t, x, \varrho_1) + \alpha_2 s_n(t, x, \varrho_2) + \ldots + \alpha_{H+1} s_n(t, x, \varrho_{H+1}) = \Phi(t, x, \varrho),$$

wo $|\varrho_j| < 1$, $j = 1, 2, \ldots, H+1$, die ϱ_j sind alle verschieden und über die Zahlen α_ν werden wir später verfügen. $\Phi(t, x, \varrho)$ hat offenbar die Form $\sum_{\nu=0}^{n} \beta_\nu \psi_\nu(x)\, \psi_\nu(t)$; der Koeffizient von $\psi_\nu(x)\, \psi_\nu(t)$ ist offenbar $\alpha_1 \varrho_1^j + \alpha_2 \varrho_2^j + \ldots + \alpha_{H+1} \varrho_{H+1}^j$, wo j wie im (89c) die Zifferanzahl von ν im dyadischen System bedeutet. Wenn wir also die Zahlen $\alpha_1, \alpha_2, \ldots, \alpha_{H+1}$ so bestimmen, daß

$$\begin{aligned} \alpha_1 + \alpha_2 + \ldots + \alpha_{H+1} &= 1 \\ \varrho_1 \alpha_1 + \varrho_2 \alpha_2 + \ldots + \varrho_{H+1} \alpha_{H+1} &= 1 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ \varrho_2^H \alpha_1 + \varrho_2^H \alpha_2 + \ldots + \varrho_{H+1}^H \alpha_{H+1} &= 1 \end{aligned}$$

gilt (was wegen $\varrho_i \neq \varrho_k$ möglich ist), dann ist wegen der Lückeneigenschaft von $f(x)$, wenn $u_n(x)$ bzw. $U_n(x)$ die n-te Partialsumme bzw. n-te Cesàro-Mittel erster Ordnung der Walsh-Reihe von $f(x)$ bedeuten,

$$u_n(x) - U_n(x) = \int_0^1 [f(t) - U_n(t)]\, \Phi(t, x, \varrho)\, dt.$$

Es ist wegen (91)

$$|u_n(x) - U_n(x)| < \left(\sum_{\nu=1}^{H+1} \frac{|\alpha_\nu|}{1 - |\varrho_\nu|} \right) \max_{0 \leqq t \leqq +1} |f(t) - U_n(t)|,$$

also ist auf Grund des bekanten Summabilitätssatzes unsere Behauptung bewiesen.

Es gilt auch die folgende Verschärfung des Satzes XVI. Es seien $N_{1j}, N_{2j}, \ldots$ diejenige ganze Zahlen, deren Zifferanzahl im dyadischen Systeme genau j ist. Dann folgt aus den Voraussetzungen des Satzes XVI auch die gleichmäßige Konvergenz im $[0, 1]$ der Teilreihen

$$\sum_{k=1}^{\infty} a_{N_{kj}} \psi_{N_{kj}}(x)$$

für $1 \leqq j \leqq H$; demzufolge ist $\sum_{k=1}^{\infty} |a_{N_{k1}}| < \infty$.

Mittels den obigen $P_k(t, x, \varrho)$ und $s_n(t, x, \varrho)$ kann man den folgenden Satz beweisen, welcher eine gewisse Ähnlichkeit mit den Sätzen des §. 2. zeigt.

Satz XVII. *Wenn $\psi_0, \psi_1, \psi_2, \ldots$ das Walsh-System und $\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots$ beliebige Zahlen mit $\sum_{\nu=0}^{\infty} \varepsilon_\nu^2 < \infty$ bedeuten, dann besitzt das System*

$$\int_0^1 f(x)\psi_{2k}(x)\,dx = \varepsilon_k$$

eine Lösung $f(x)$, deren Walsh-Entwicklung überall gleichmäßig konvergiert.

Beweis. Nach STEINHAUS und KACZMARZ[25]) existiert eine überall stetige $g(x)$ mit $g(x) \sim \sum_{\nu=0}^{\infty} b_\nu \psi_\nu(x)$, $b_{2k} = 2\varepsilon_k$ $(k = 1, 2, \ldots)$. Es sei

$$f_k(x) = \int_0^1 g(t) P_k\left(t, x, \frac{1}{2}\right) dt. \tag{93}$$

$f_k(x)$ ist offenbar ein Abschnitt von $f_{k+1}(x)$, also nach (89b) ist für $n \geqq k$

$$\int_0^1 f_n(t)\psi_{2k}(t)\,dt = \varepsilon_k.$$

Wenn $v_n^{(k)}(x)$ und $V_n^{(k)}(x)$ die n-te Partialsumme bzw. das n-te Cesàro-Mittel erster Ordnung von $f_k(x)$, $W_n(x)$ das n-te Cesàro-Mittel von $g(x)$ bedeuten, dann ist wegen (92) und

$$V_n^{(k)}(x) - V_m^{(k)}(x) = \int_0^1 [W_n(t) - W_m(t)]\, s_n\left(t, x, \frac{1}{2}\right) dt, \quad m > n$$

die Funktion $\lim_{k \to \infty} f_k(x) = f(x)$ überall stetig. Wegen

25) H. STEINHAUS—S. KACZMARZ, Le système orthogonal de M. Rademacher, *Studia Math.*, 2 (1930), S. 231—247.

$$v_n^{(k)}(x) - V_n^{(k)}(x) = \int_0^1 [g(t) - W_n(t)]\, s_n\left(t, x, \frac{1}{2}\right) dt$$

ist also die Folge $v_n^{(k)}(x)$ gleichmäßig konvergent, also offenbar auch die Partialsummen von $f(x)$, qu. e. d.

Ein zum Satz XVII analoger Satz habe ich für Potenzreihen in meiner Arbeit[23]) ausgesprochen; dies gilt aber nur, wenn wir die Glieder in gewisser Art umordnen. Dem Satz XV entsprechend gilt aber, wie es sich analogerweise beweisen läßt, der

Satz XVIII. *Die Fourierreihe* $\sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx)$ *der überall stetigen* $f(x)$ *habe die Eigenschaft, daß es eine ganze Zahl* H *gibt so, daß alle Koeffizienten* a_n, b_n *verschwinden, deren Index im dyadischen System wenigstens* $(H+1)$*-ziffrig ist. Dann konvergiert die Reihe in jedem Intervalle gleichmäßig.*

Für $H = 1$ hat diesen Satz schon KOLMOGOROFF[26]) bewiesen; in diesem Falle gilt auch absolute Konvergenz.

Anhang I.

Herr KACZMARZ[17]) bewies den Satz: *die notwendige und hinreichende Bedingung dafür, daß die Folge der Konstanten* $\alpha_0, \alpha_1, \alpha_2, \ldots$ *die Eigenschaft habe, eine beliebige Funktion der Klasse* L, $f(x) \sim \sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx)$ *in eine solche der Klasse* L_p, $p > 1$, $g(x) \sim \sum_{n=0}^{\infty} \alpha_n (a_n \cos nx + b_n \sin nx)$ *zu überführen, ist, daß* $h(x) \sim \sum_{n=0}^{\infty} \alpha_n \cos nx$ *zur Klasse* L_p *gehöre.*

In den folgenden geben wir für diesen Satz einen, von dem Kaczmarzschen verschiedenen, direkten Beweis.

Bezeichnet $S_n(x)$, $\sigma_n(x)$ das n-te Cesàro-Mittel erster Ordnung der Fourierreihe von $g(x)$ bzw. $h(x)$, so gilt

$$S_n(x) = \frac{1}{2\pi} \int_0^{2\pi} \sigma_n(x+t) f(t)\, dt,$$

---

26) A. KOLMOGOROFF, Une contribution à l'étude de la convergence des séries de Fourier, *Fundamenta Math.*, 5 (1924) S. 96—97.

also nach dem bekanntem Youngschen Satze[27])

$$\int_0^{2\pi} |S_n(x)|^p\,dx \leqq \int_0^{2\pi} |\sigma_n(t)|^p\,dt \int_0^{2\pi} |f(t)|\,dt,$$

woraus die Hinlänglichkeit der Bedingung folgt.

Es sei nun $a_0, a_1, a_2, \ldots$ eine positive konvexe Nullfolge. Es ist leicht einzusehen, daß dann die Folge $-\frac{1}{a_0}, -\frac{1}{a_1}, -\frac{1}{a_2}, \ldots$ eine konvexe Folge darstellt, welche gegen $-\infty$ strebt. Dann ist die Reihe $\sum_{\nu=0}^{\infty} a_\nu \cos\nu x$ bekanntlich die Fourierreihe einer Funktion $F(x)$ aus der Klasse L[28]). Wir bestimmen die Zahlen $\beta_0, \beta_1, \ldots$ so, daß

$$\sigma_n(x) = \sum_{k=0}^{n} \beta_k S_k(x). \tag{94}$$

Da

$$\sigma_n(x) = \sum_{\nu=0}^{n} \frac{n-\nu+1}{n+1} a_\nu \cos\nu x, \quad S_k(x) = \sum_{\mu=0}^{k} \frac{k-\mu+1}{k+1} \alpha_\mu a_\mu \cos\mu x$$

ist, gilt wegen (94)

$$\frac{n-l+1}{n+1} a_l = \frac{1}{l+1} \alpha_l a_l \beta_l + \frac{2}{l+2} \alpha_l a_l \beta_{l+1} + \ldots + \frac{n-l+1}{n+1} a_l \alpha_l \beta_n,$$

$$\left\{\begin{aligned} \beta_l &= \frac{l+1}{n+1}\left(\frac{n-l+1}{a_l} - \frac{2n-2l}{a_{l+1}} + \frac{n-l+1}{a_{l+2}}\right), \\ &\qquad (l=0, 1, \ldots, n-2) \\ \beta_{n-1} &= \frac{2n}{n+1}\left(\frac{1}{a_{n-1}} - \frac{1}{a_n}\right), \\ \beta_n &= \frac{1}{a_n}. \end{aligned}\right. \tag{95}$$

Da die Folge $-\frac{1}{a_k}$ konvex ist, folgt aus (95)

$$\sum_{l=0}^{n} |\beta_l| = \frac{1}{a_n} - \frac{1}{a_0}. \tag{96}$$

Aus (94) und (96) folgt, daß

---

27) Siehe z. B. Zygmund, *Trigonometrical Series*, S. 71. Anwendung mit $q=1$.

28) Satz von Young. Siehe Zygmund, *Trigonometrical Series*, S. 109.

$$I_n = \left(\int_0^{2\pi} |\sigma_n(x)|^p \, dx \right)^{1/p} \leqq \frac{2}{a_n} \max_{k \leqq n} \left(\int_0^{2\pi} |S_k(x)|^p \, dx \right)^{1/p}.$$

Wenn die linke Seite unbeschränkt wäre, dann könnte man die Folge $a_0, a_1, a_2, \ldots$ so wählen, daß auch die Folge $a_n I_n$ unbeschränkt sei. Dann könnte aber diese $g(x)$ wegen $\overline{\lim\limits_{k \to \infty}} \left(\int_0^{2\pi} |S_k(x)| \, dx \right)^{1/p} = \infty$ nicht zu L_p gehören, qu. e. d.

Anhang II.

Ich nenne hier eine Funktionenreihe $\sum\limits_{n=0}^{\infty} f_n(x)$ eine B-Reihe bezüglich des Intervalls $a < x < b$, wenn ihre Abelschen Summen für $a < x < b$ gleichmäßig beschränkt sind, eine C-Reihe, wenn ihre Abelschen Summen für $a < x < b$ gleichmäßig beschränkt und konvergent sind; eine Folge $\alpha_0, \alpha_1, \ldots, \alpha_n, \ldots$ mit konstanten Gliedern quasi-vollmonoton, wenn $\Re(\alpha_0), \Re(\alpha_1), \ldots, \Re(\alpha_n), \ldots$ und $\Im(\alpha_0), \Im(\alpha_1), \ldots, \Im(\alpha_n), \ldots$ die Differenz zweier beschränkter vollmonotoner Folgen sind. Es gelten die Sätze:

Satz A. *Dafür, daß eine numerische Faktorenfolge α_0, $\alpha_1, \ldots, \alpha_n, \ldots$ jede B-Reihe bezüglich des Intervalls $a < x < b$ in eine ebensolche überführe, ist notwendig und hinreichend, daß sie quasi-vollmonoton sei.*

Satz B. *Eine jede C-Reihe bezüglich des Intervalls $a < x < b$ in eine ebensolche überführende, numerische Faktorenfolge*[29]) *ist quasi-vollmonoton.*

Beweis des Satzes A. Ist die Folge $\alpha_0, \alpha_1, \ldots, \alpha_n, \ldots$ quasi-vollmonoton, so gibt es eine Darstellung $\alpha_n = \int_0^1 t^n dp(t)$ $(n = 0, 1, 2, \ldots)$ mit $\int_0^1 |dp(t)| < \infty$[30]). Gilt für die numerische Reihe $\sum\limits_{n=0}^{\infty} u_n$ die Ungleichung $\left| \sum\limits_{n=0}^{\infty} u_n r^n \right| < M$, $M > 0$ uud fest, für

[29]) Die Existenz solcher nichttrivialer Folgen folgt aus Hilfssatz II dieses Anhanges.

[30]) F. Hausdorff, Summationsmethoden und Momentfolgen, *Math. Zeitschrift*, 9 (1921), S. 74–109.

$0 < r < 1$, so ist

$$\underset{0<r<1}{\text{Max}} \left| \sum_{n=0}^{\infty} \alpha_n u_n r^n \right| = \underset{0<r<1}{\text{Max}} \left| \int_0^1 \sum_{n=0}^{\infty} u_n (rt)^n \, dp(t) \right| < M \int_0^1 |dp(t)|.$$

Überführt die Faktorenfolge $\alpha_0, \alpha_1, \ldots, \alpha_n, \ldots$ jede B-Reihe in eine B-Reihe, so gilt für ein beliebiges Polynom $P(z) = \sum_{n=0}^{N} a_n z^n$

$$\left| \sum_{n=0}^{N} a_n \alpha_n \right| < C \underset{0<z<1}{\text{Max}} |P(z)| \tag{1}$$

mit nur von den α abhängigem C.[31]) Nach einem wohlbekannten Satze des Herrn F. Riesz[32]) folgt daraus die Erfüllbarkeit des Gleichungssystems

$$\int dp(t) = \alpha_0, \int t \, dp(t) = \alpha_1, \ldots, \int t^n dp(t) = \alpha_n, \ldots$$

durch eine Funktion $p(t)$ mit $\int_0^1 |dp(t)| < \infty$ d. h. die Quasi-Vollmonotonität der Folge $\alpha_0, \alpha_1, \ldots, \alpha_n \ldots$

Beim Beweise des Satzes B stütze ich mich auf folgende Hilfssätze:

Hilfssatz I. *Ist die numerische Folge $\alpha_0, \alpha_1, \ldots, \alpha_n, \ldots$ nicht quasi-vollmonoton, so gibt es eine B-Reihe $\sum_{n=0}^{\infty} f_n(x)$ bezüglich des Intervalls $a < x < b$, für welche $\overline{\lim} \left| \sum_{n=0}^{\infty} \alpha_n f_n(x) (r_2^n - r_1^n) \right| = \infty$ für $a < x < b$, $0 < r_1 < r_2 < 1$, $\lim r_1 = \lim r_2 = 1$ gilt.*

[31]) Wäre (1) nicht erfüllt, so gäbe es eine Folge, von Polynomen $P_1(z), P_2(z), \ldots, P_k(z) = \sum_{n=0}^{N_k} a_{nk} z^n, \ldots$ mit $\underset{0<z<1}{\text{Max}} |P_k(z)| < 1$ und

$$\lim_{k=\infty} \left| \sum_{n=0}^{N_k} a_{nk} \alpha_k \right| = \infty.$$

Für die Funktionenreihe $\sum_{n=0}^{\infty} f_n(x)$, wo $f_n(x_k) = a_{nk}$ für $x = x_k$, $f_n(x) = 0$ für $x \neq x_k$, $a < x_k < b$, $a_{nk} = 0$ für $n > N_k$ gilt also

$$\left| \sum_{n=0}^{\infty} f_n(x) \right| < 1, \ \overline{\lim} \left| \sum_{n=0}^{\infty} \alpha_n f_n(x) r^n \right| = \infty$$

für $a < x < b$, $0 < r < 1$.

[32]) F. Riesz, Sur certains systémes singuliers d'équations integrales, *Annales de l'École Normale Supérieure*, 28 (1911), S. 33–62.

Beweis. Aus $\overline{\lim}\left|\sum_{n=0}^{\infty} \alpha_n f_n(x)(r_2^n - r_1^n)\right| < \infty$ für $a < x < b$, $0 < r_1 < r_2 < 1$, $\lim r_1 = 1$ für jede B-Reihe bezüglich eines Intervalls $a < x < b$ folgt für ein beliebiges Polynom $P(z) = \sum_{n=0}^{N} a_n z^n$

$$(2) \qquad \left|\sum_{n=0}^{N} a_n \alpha_n (1 - r^n)\right| < C \max_{0<z<1} |P(z)|$$

für jedes die Bedingung $0 < r_1 < r < 1$, r_1 fest, erfüllende r mit nur von den α abhängigem C.[38]) Die Folge $\alpha_0, \alpha_1(1-r_1), \ldots, \alpha_n(1-r_1^n), \ldots$ und daher auch $\alpha_0, \alpha_1, \ldots, \alpha_n, \ldots$ ist also quasi-vollmonoton.

Hilfssatz II. *Eine numerische Faktorenfolge*

$$\beta_0, \beta_1 = \int_0^1 t\, dp(t), \ldots, \beta_n = \int_0^1 t^n\, dp(t), \ldots$$

mit totalsetiger Belegungsfunktion überführt jede B-Reihe bezüglich $a < x < b$ in eine C-Reihe bezüglich $a < x < b$.

Beweis. Ist $\sum_{n=0}^{\infty} u_n$ eine numerische Reihe mit $\left|\sum_{n=0}^{\infty} u_n r^n\right| < M$, $M > 0$ und fest für $0 < r < 1$, so gilt für $0 < r_1 < r_2 < 1$, wenn $\varepsilon > 0$, sonst beliebig, das Polynom $P(z) = \sum_{m=0}^{M} a_m z^m$ die Ungleichung $\int_0^1 |d[p(t) - P(t)]| < \frac{\varepsilon}{4M}$ erfüllt, $\sum_{n=0}^{\infty} u_n z^n = g(z)$ für $0 < z < 1$, $\int_0^1 t^n P(t)\, dt = \beta_n'$ gesetzt wird,

$$\left|\sum_{n=0}^{\infty} u_n(\beta_n - \beta_n')(r_2^n - r_1^n)\right| = \left|\int_0^1 [g(tr_1) - g(tr_2)]\, d[p(t) - P(t)]'\right| < \frac{\varepsilon}{2},$$

$$\left|\sum_{n=0}^{\infty} u_n \beta_n'(r_2^n - r_1^n)\right| =$$

$$= \left|\sum_{m=1}^{M} m a_m \left[\frac{1}{r_1^m}\int_{r_1}^{r_2} r^{m-1} g(r)\, dr + \left(\frac{1}{r_1^m} - \frac{1}{r_2^m}\right)\int_0^{r_2} r^{m-1} g(r)\, dr\right]\right| <$$

$$< (r_2 - r_1)\, Q$$

---

[38]) (2) ergibt sich analog wie (1).

mit nur von M und den a abhängigem Q. Es ist also für $r_2 - r_1 < \frac{1}{2Q}$

$$\left| \sum_{n=0}^{\infty} \beta_n u_n (r_2^n - r_1^n) \right| < \varepsilon.$$

H i l f s s a t z III. *Jede Funktionenreihe*[34]) $\sum_{n=0}^{\infty} f_n(x)$ *mit* $\overline{\lim} \left| \sum_{n=0}^{\infty} f_n(x)(r_2^n - r_1^n) \right| = \infty$ *für* $a < x < b$, $0 < r_1 < r_2 < 1$, $\lim r_1 = 1$ *kann durch eine Faktorenfolge* $\beta_0, \beta_1 = \int_0^1 t\,dp(t), \ldots, \beta_n = \int_0^1 t^n\,dp(t), \ldots$ *mit totalstetiger Belegungsfunktion in eine Reihe von ebensolcher Eigenschaft überführt werden.*

B e w e i s. Ohne Beschränkung der Allgemeinheit kann vorausgesetzt werden, daß es Folgen $x_1, x_2, \ldots, x_k, \ldots$; $r_{11}, r_{12}, \ldots, r_{1k}, \ldots$; $r_{21}, r_{22}, \ldots, r_{2k}, \ldots$ mit $a < x_k < b$, $0 < r_{1k} < r_{2k} < 1$, $\lim r_{1k} = 1$ und $\lim \sum_{n=0}^{\infty} f_n(x_k)(r_{2k}^n - r_{1k}^n) = \infty$ gibt. Ist $\sum_{n=0}^{\infty} f_n(x_k)[(tr_{2k})^n - (tr_{1k})^n] > \frac{1}{2} \sum_{n=0}^{\infty} f_n(x_k)(r_{2k}^n - r_{2k}^n)$ für $t_k < t < 1$, $p_k(t) = 0$ für $0 < t < t_k$, $p_k(t) = \frac{t - t_k}{1 - t_k}$ für $t_k < t < 1$, so läßt sich nach dem aus der Theorie der singulären Integrale wohlbekannten Lebesgueschen Verfahren die Funktion $p(t) = \sum_{l=1}^{\infty} \gamma_l p_{k_l}(t)$ mit konstanten γ und $\sum_{l=1}^{\infty} |\gamma_l| < \infty$ so konstruieren, daß die zugehörige Folge

$$\beta_0, \beta_1 = \int_0^1 t\,dp(t), \ldots, \beta_n = \int_0^1 t^n\,dp(t), \ldots$$

von der gewünschten Beschaffenheit sei.

B e w e i s d e s S a t z e s B. Ist die numerische Folge α_0, $\alpha_1, \ldots, \alpha_n, \ldots$ nicht quasi-vollmonoton, so gibt es eine B-Reihe bezüglich des Intervalls $a < x < b$, $\sum_{n=0}^{\infty} f_n(x)$, mit

$$\overline{\lim} \left| \sum_{n=0}^{\infty} \alpha_n f_n(x)(r_2^n - r_1^n) \right| = \infty$$

---

[34]) Hierbei ist vorausgesetzt, daß die Potenzreihen $\sum_{n=0}^{\infty} f_n(x) z^n$ für $|z| < 1$ konvergieren.

für $a < x < b$, $0 < r_1 < r_2 < 1$, $\lim r_1 = \lim r_2 = 1$. Ist

$$\beta_0, \beta_1 = \int_0^1 t\, dp(t), \ldots, \beta_n = \int_0^1 t^n\, dp(t), \ldots$$

die nach Hilfssatz III existierende Folge mit totalstetiger Belegungsfunktion, für welche $\overline{\lim} \left| \sum_{n=0}^{\infty} \alpha_n \beta_n f_n(x) (r_2^n - r_1^n) \right| = \infty$ für $a < x < b$, $0 < r_1 < r_2 < 1$, $\lim r_1 = \lim r_2 = 1$, so wird die C-Reihe $\sum_{n=0}^{\infty} \beta_n f_n(x)$ bezüglich des Intervalls $a < x < b$ durch $\alpha_0, \alpha_1, \ldots, \alpha_n, \ldots$ in keine ebensolche übergeführt.

(Eingegangen von 5. Juni 1937 bis 13. Juni 1940;
umgearbeitet eingegangen am 31. Dezember 1941.)

Bibliographie.

E. A. Weiss, Einführung in die Liniengeometrie und Kinematik (Teubners math. Leitfäden, Bd. 41), VI + 122 S., Leipzig und Berlin, B. G. Teubner, 1935.

Das zu der selben Sammlung gehörige Bändchen von L. BIEBERBACH, *Einleitung in die höhere Geometrie* enthält die Elemente der Liniengeometrie. Der Verfasser wendet sich an Leser, die mit den Grundbegriffen einigermaßen schon vertraut sind und führt sie auf einen bis zum Schluß reizvollen Weg in die tiefsten Gebiete der Studyschen Gedankenwelt ein. Das Büchlein umfaßt ein überaus großes Gebiet der algebraischen (nicht differentiellen) Liniengeometrie und Kinematik. In der Behandlung kommt eine einheitliche Betrachtungsweise zur Geltung: „Liniengeometrie ist nach FELIX KLEIN die Geometrie einer quadratischen Mannigfaltigkeit in einem Raume von 5 Dimensionen, Kinematik — die Geometrie mit der Bewegung als Raumelement — nach EDUARD STUDY die Geometrie einer quadratischen Mannigfaltigkeit in einem Raume von 7 Dimensionen und als solche eine natürliche Verallgemeinerung der Liniengeometrie" (Vorwort). Diese Betrachtungsweise, sowie die zahlreichen Anwendungen erleichtern das Eindringen in dieses schöne, im übrigen aber keineswegs leicht zugängliche Gebiet der Geometrie.

Die Hervorhebung nur einiger Stichwörter erhellt zur genüge den reichhaltigen Inhalt des Büchleins. Es seien mit besonderer Rücksicht auf die Anwendungen genannt: Weitzenböcksche Ketten, Kummersche Konfiguration, Geraden-Kugel-Transformation (nach STUDY), Dupinsche Zykliden, Studysche Doppelfünf, Cliffordsche Parallelen, Plückersches Zylindroid, Bewegungen und Umlegungen (mit Hilfe der Hamiltonschen Quaternionen).

L. Fejes.

H. Ertel, Methoden und Probleme der dynamischen Meteorologie (Ergebnisse der Math. und ihrer Grenzgebiete, fünfter Band, Heft 3), IV + 122 S., Berlin, J. Springer, 1938.

Das Büchlein bietet eine Übersicht über ausgewählte Probleme der dynamischen Meteorologie und über die Methoden, die zu deren Lösung

Anwendung finden. Es ist kein Lehrbuch, vielmehr ist es dazu geeignet, den schon in mancher Hinsicht überholten Enzyklopädieartikel von EXNER und TRABERT zu ersetzen. Alle wichtigen Problemkreise werden gestreift, viele werden, besonders die, in denen der Verfasser manche schöne, mitunter grundlegende Resultate erzielt hat, eingehend dargestellt. Die Vektorenrechnung wird überall angewendet, und zwar in koordinatenmäßiger Schreibweise. In der Einleitung gibt Verfasser eine klare, exakte Definition der dynamischen Meteorologie und deren Aufgaben. Die dabei entwickelten Ansichten sind vielfach sehr beachtenswert.

Der Inhalt des Werkes gestaltet sich folgendermaßen. In einem ersten Kapitel werden die nötigen thermo-hydrodynamischen Grundlagen und Vorkenntnisse gegeben. Das zweite Kapitel enthält dann die allgemeine Dynamik der Atmosphäre. Die hydrodynamischen Gleichungen werden in rotierendem Koordinatensystem abgeleitet, die Begriffe: ausgeglichene Bewegung, Turbulenz, Geopotential und verschiedene Grundbegriffe der Bjerknesschen physikalischen Hydrodynamik eingeführt. Dann folgt eine überaus originelle Behandlung des Zirkulationstheorems von BJERKNES. Anschließend werden die Gesetze der atmosphärischen Dynamik auf ein allgemeines Variationsprinzip zurückgeführt. Dann werden die atmosphärischen Energieumsetzungen behandelt: ein Problemkreis, der ohne Zweifel zentrale Bedeutung besitzt. Dabei werden unter anderen besonders die Arbeiten von MARGULES, REFSDAL und HORIGUTI berücksichtigt.

Im dritten Kapitel werden einzelne Problemstellungen besprochen. Nach der Behandlung der Stabilitäts- und Labilitätsfragen wird das Problem der mittleren Temperaturverteilung der Atmosphäre, also des Zustandekommens der Tropo- und Stratosphäre ausführlich dargestellt. Bei der Behandlung der Zustandsänderungen in vertikalen Luftsäulen und der Theorie der Advektion werden auch die diesbezüglichen, in gewisser Richtung bahnbrechenden Arbeiten des ungarischen Forschers L. STEINER hervorgehoben. Nach der Behandlung der Eigenschaften stationärer Windfelder werden die Diskontinuitäten besprochen: dabei kommt auch die Gleichgewichtsbedingung für das Auftreten eines Tropopausentrichters nach PALMÉN zur Ableitung. Nach einer Ableitung der atmosphärischen Störungsgleichungen werden die Grundbedingungen formuliert, denen eine Theorie der Zyklonenbildung, also eine Lösung der Grundaufgabe der Meteorologie genügen muß. Zum Schluß wird eine (ebenfalls originelle) Relaxationstheorie der atmosphärischen Gradientwindabweichungen gegeben und dadurch manche scheinbar widersprechende, theoretisch und empirisch gewonnene frühere Resultate in Einklang gebracht.

Eine überaus reiche Literaturliste (246 Angaben) ergänzt das treffliche Buch, welches seine Zielsetzung vollkommen erreicht und in der heute in äußerstem Maße angeschwollenen meteorologischen Literatur eine hervorragende Stellung einnimmt.

G. Tóth.

Wilhelm Blaschke und Gerrit Bol, Geometrie der Gewebe, Topologische Fragen der Differentialgeometrie (Grundlehren der math. Wissenschaften, Bd 49), VIII + 339 S., Berlin, J. Springer, 1938.

Gegenstand des Buches ist die Erforschung der topologischen Eigenschaften, die gewisse als „Gewebe" bezeichnete Kurvenscharen besitzen. Das Buch zerfällt in drei Abschnitte. Im ersten werden die einfachsten Gewebe behandelt und zwar von rein topologischem Standpunkt, d. h. es werden hier keinerlei Differenzierbarkeitsannahmen gemacht. Im Vordergrund der Betrachtungen stehen die sogenannten Schließungsfragen. Betrachtet man z. B. ein Kurven-3-Gewebe, so kann man dasselbe dann und nur dann auf ein „regelmäßiges" Kurven-3-Gewebe topologisch abbilden, wenn sich sämtliche Brianchonschen Sechsecke schließen. Dabei besteht ein regelmäßiges Kurven-3-Gewebe aus drei Scharen von parallelen Geraden, die eine Winkel von 60° einschließen. Solche Kurven-3-Gewebe werden als Sechseckgewebe bezeichnet. Es werden die Grundlagen für einen axiomatischen Aufbau der Theorie der Gewebe gegeben, die sich auf die Gruppentheorie stützen. Die Bedeutung der Schließungsfragen zeigt sich hier von neuem. Die Untersuchungen werden dann auch auf den Raum ausgedehnt. Sowohl bei ebenen als auch bei räumlichen Geweben handelt es sich um folgende fundamentale Frage: wann läßt sich ein n-Gewebe auf n Geradenbüschel abbilden. Ein Kennzeichen dafür besteht darin, daß jedes in dem n-Gewebe enthaltene 3-Gewebe ein Sechseckgewebe bildet. Im zweiten Abschnitt ist die zugrunde gelegte Gruppe nicht mehr die aller topologischen Abbildungen, sondern die aller hinreichend oft differenzierbaren. Gegenstand dieses Abschnittes bildet die invariantentheoretische Charakterisierung der Gewebe. Benutzt werden dabei diejenigen linearen Differentiatoren, aus denen die infinitesimalen Transformationen in der Lieschen Theorie aufgebaut werden. Um etwa ein Kurven-3-Gewebe zu charakterisieren, geht man so vor: jede Kurvenschar wird durch einen Differentiator dargestellt. Die topologischen Invarianten der drei Differentiatoren, die bei gewissen Umnormierungen invariant sind, liefern dann sämtliche Invarianten des Gewebes.

Die Probleme des dritten Abschnittes gruppieren sich um das „Abelsche Theorem in der Gewebegeometrie". Es handelt sich um eine Charakterisierung von algebraischen Gebilden. Der Hauptsatz ist dabei folgender. Eine Schar von n stetig-differenzierbaren Kurvenbögen bestimmen dann eine algebraische Kurve n-ter Ordnung, wenn es für jede Schar einen Parameter t gibt, so daß $\sum_{i=1}^{n} t_i = k$ gilt für n Punkte einer Geraden, die auf je einer Kurve liegen. Faßt man die Frage rein funktionentheoretisch auf, so wird man zu Abels Theorem geführt. Die Abelschen Parameter t_i werden im weiteren benutzt, um den Höchstrang von Geweben zu charakterisieren. Die Betrachtung der Flächengewebe von höchstem Rang führt dann zu Gedankengängen, die der Weylschen Geometrie angehören.

Die zahlreichen Arbeiten der Blaschkeschen Schule finden hier eine zu begrüßende einheitliche Verschmelzung.

O. Varga.

Eliakim Hastings Moore, General Analysis, Part II: The Fundamental Notions of General Analysis, with the cooperation of RAYMOND WALTER BARNARD (Memoirs of the American Philosophical Society held at Philadelphia for promoting useful knowledge, Volume I, Part II), VI + 255 pages, Philadelphia, The American Philosophical Society, 1939.

The present second volume of Moore's General Analysis starts with a chapter on the general limit notion introduced in 1922 by MOORE and SMITH, and which has found since important applications not only in Moore's own investigations, but also in other domains of recent mathematical research. Let us recall its definition. Let A be the system of real or complex numbers or quaternions. Let L be an arbitrary semi-ordered set, i. e. for the elements l of which a transitive relation $l_1 > l_2$ is defined such that to each pair l_1, l_2 there exists an l with the property $l > l_1$, $l > l_2$. If $a(l)$ is a function on L with values in A, then a number a_0 in A is called the limit of $a(l)$, if to any $\varepsilon > 0$ there corresponds an $l_0(\varepsilon)$ such that $|a(l) - a_0| \leqq \varepsilon$ for $l > l_0(\varepsilon)$. — The following chapters develop the proper fundamental notions of General Analysis, such as the modular space, the generalised integral of MOORE, and their applications to the theory of modular matrices and linear continuous transformations. Each chapter is headed by an introduction describing in relatively non-technical language the contents of the chapter.

Béla de Sz. Nagy.

E. A. Weiss, Punktreihengeometrie, VIII + 232 S., Leipzig und Berlin, B. G. Teubner, 1939.

Punktreihe bezeichnet eine mit Parameter versehene Gerade. Punktreihengeometrie bedeutet die Geometrie, deren Raumelemente Punktreihen sind. (Punktbüschel bedeutet hier eine Punktreihe mit linearer Parameterdarstellung.)

Das vorliegende Buch bezweckt eine Einführung in die projektive Geometrie mehrdimensionaler Räume. Dies wird hauptsächlich durch die Abbildung der Ebene und des Raumes auf die Punkte des 5- bzw. 7-dimensionalen projektiven Raumes R_5 bzw. R_7 erreicht. In den verschiedensten Problemen dieses Werkes spielen die Segreschen Mannigfaltigkeiten eine wichtige Rolle.

Das Buch gliedert sich in 4 Kapitel.

Kapitel I gibt eine Einführung in die symbolische Bezeichnungsweise und behandelt die Geometrie der Punktreihen *auf der Geraden* und ihre

Abbildung auf den Raum R_3. Bei dieser Abbildung werden die singulären Punktreihen auf eine Regelfläche zweiter Ordnung abgebildet.

Kapitel II beschäftigt sich mit der Punktreihengeometrie *in der Ebene*. Es werden hier unter anderem die Übertragungsprinzipien von CLEBSCH und HESSE, die Segreschen Mannigfaltigkeiten, die kubischen Regelflächen, die Theorie der apolaren Kegelschnitte, der Ponceletsche Schließungssatz und das Nullsystem im R_3 eingehend behandelt. Die hier auftretende dreidimensionale Segresche Mannigfaltigkeit von R_5 und ihre automorphe Kollineationen und Korrelationen werden ausführlich be sprochen.

Kapitel III untersucht die Funktreihengeometrie *im Raume* und ihre Abbildung auf die Punkte von R_7. Unter den überaus reichen Anwendungen erwähnen wir die folgenden: Räumliche Analoga der Übertragungsprinzipien von CLEBSCH und HESSE, die Regelfläche und die rationale Raumkurve vierter Ordnung, die Verallgemeinerung des Desarguesschen Satzes im R_n, das Trialitätsprinzip, die Engelsche Ebenengeometrie und die euklidische Geraden-Kugel-Transformation.

Kapitel IV beschäftigt sich mit der trilinearen Verwandtschaft zwischen drei Punktreihen im R_2, R_3, R_4 und R_5, mit der Abbildung der trilinearen Verwandschaften auf die Punkte von R_7 und mit der Begründung der nichteuklidischen Punktreihengeometrie.

Dieses Buch gibt nicht nur eine anschauliche Einführung in die mehrdimensionale projektive Geometrie, es liefert auch neue geometrische Gesichtspunkte und ein reichhaltiges Material an verschiedenartigen geometrischen Problemen, und somit auch gute Anregungen zu neuen Forschungen. Dieses Werk wird gewiß dazu beitragen, die Tradition der projektiven Geometrie wachzuhalten und in großen Kreisen zu verbreiten.

Gy. v. Sz. Nagy.

Enzyklopädie der mathematischen Wissenschaften mit Einschluß ihrer Anwendungen, herausgegeben im Auftrage der Akademien der Wissenschaften zu Berlin, Göttingen, Heidelberg, Leipzig, München und Wien sowie unter Mitwirkung zahlreicher Fachgenossen, Leipzig und Berlin, B. G. Teubner.

Es ist natürlich zu fragen, wozu eine Neubearbeitung der wohlbewährten, den zu Beginn gestellten Zwecken vorzüglich entsprechenden Enzyklopädie dient.

Als die ersten Hefte der alten Encyklopädie erschienen, hat man allgemein die Auffassung vertreten, die Wahrheiten der Mathematik seien unabhängig von jeder menschlicher Denktätigkeit. Nach dieser, zuweilen als logistisch bezeichneten Auffassung verliert eine einmal festgestellte Wahrheit seinen Wert nie; die Entwicklung der Mathematik besteht darin, daß neben den schon bewiesenen Theoremen weitere entdeckt werden, die eigentlich schon früher da waren, aber bis dahin unbekannt geblieben sind. Dieser Auffassung entsprechend bestand die Aufgabe der alten Encyklopädie darin, möglichst alle bewiesene mathematische Sätze systematisch zu registrieren; der Entwicklung der Wissenschaft wurde durch Herausgabe von Ergänzungsheften über die neueren Ergebnisse der einzelnen Zweige Rechnung getragen.

Daß nun statt weiteren Ergänzungsbänden eine Neufassung des Ganzen unternommen wird, hat vor allem den Grund, daß unsere Auffassung über die Mathematik im Verändern ist. Psychologistische Gesichtspunkte, die allerdings schon früher für einzelne Mathematiker, wenn auch oft unbewußt und meistens unausgesprochen, maßgebend waren, treten mehr und mehr in den Vordergrund. Einige dieser Gesichtspunkte betreffen die Beziehung des Mathematikers zu seiner Wissenschaft, so z. B. unterscheidet man immer schärfer zwischen interessanten und weniger interessanten Ergebnissen; man legt einer effektiven Konstruktion einen größeren Wert bei, als einem rein logischen Existenzbeweis. Es ist nur ein Schritt, hiervon zur Beziehung des Mathematikers zu seinem Mitmenschen (Leser, Studierende) zu gelangen: pädagogische Gesichtspunkte werden untrennbar von der mathematischen Forschung; man legt einen großen Gewicht auf die Vereinfachung von Beweisen und auf die Heuristik; statt des alten „was" und „wie" fragt man immer öfter „warum" und „wozu". Diese Auffassung der Wissenschaft erfordert nicht nur eine einmalige, sondern eine stete Neuformulierung unserer Behauptungen.

In welchem Maße diesen Gesichtspunkten die Neubearbeitung der Enzyklopädie entsprechen wird, läßt sich nicht im Voraus sagen; jedenfalls begrüßen wir, daß bei der Verarbeitung der Literatur nicht eine bedingungslose Vollständigkeit erstrebt wird, weil das zu einer Gleichstellung allgemeiner Ergebnisse von überragender Bedeutung mit unwichtigen Einzelergebnissen führen würde; ferner, daß die Beweise der wichtigsten Sätze vielfach in den Hauptpunkten etwa so weit dargestellt werden, daß sich ein Fachmann daraus den ganzen Beweis selbst aufbauen kann.

L. Kalmár.

Band I: Algebra und Zahlentheorie; I. Teil: A. Grundlagen, B. Algebra, zweite völlig neubearbeitete Auflage, herausgegeben von H. HASSE und E. HECKE.

Heft 2, ausgegeben am 28. Juni 1939.

3. FRIEDRICH BACHMANN, *Aufbau des Zahlensystems,* 28 S.

Es wird über die axiomatische Theorie der natürlichen Zahlen nebst den Rechenoperationen, Anordnung und Verwendung der Zahlen als Ordinal- und Kardinalzahlen, über die mengentheoretische Modellbildung für die Peanoschen Axiome, ferner über die Einführung der negativen und der rationalen Zahlen, der reellen Zahlen (durch Fundamentalfolgen, durch Schnitte und durch andere Methoden) und schließlich der komplexen Zahlen berichtet.

4. KONRAD KNOPP, *Darstellung der reellen Zahlen durch Grenzprozesse.* 30 S.

Es werden die Prozesse besprochen, die zur Darstellung von reellen (insbesondere irrationalen) Zahlen dienen: Schnitt (insbesondere untere, obere Grenze, unterer, oberer Limes), Schachtelung (Halbierung, Systembruch), und Folgen (unendliche Reihen, Produkte, Kettenbrücne). Als besondere Darstellungsformen werden die Cantorschen Reihen, verschiedene Stammbruchentwicklungen und das Cantorsche Produkt angeführt. Schließlich wird über die Methoden zur Beurteilung der Güte der Approximation und über die Anwendungen auf Irrationalitätsfragen berichtet.

5. E. KAMKE, *Allgemeine Mengenlehre*, 56 S.

Es wird über die allgemeine Mengenrechnung, über die Äquivalenztheorie und über Kardinalzahlen, ferner über die Ordnungstheorie und über Ordnungstypen und Ordnungszahlen berichtet.

Bei der Verarbeitung eines umgestrittenen Gebietes erwartet man von einem Enzyklopädieartikel eine objektive Nebeneinanderstellung der verschiedenen Auffassungen; eine Unterordnung aller üblichen Auffassungen einer vom Verfasser vertretenen Standpunkt, wie man es hier vom Beginn an, insbesondere auch im Kapitel: Kritische Bemerkungen über die Grundlagen der Mengenlehre, findet, entspricht eher der Natur eines Lehrbuches. Der eingenommene Standpunkt wird als eine bereinigte Form der naiven Mengenlehre in Anlehnung an A. N. WHITEHEAD und B. RUSSELL bezeichnet; daß es sich nicht um den Standpunkt der Typentheorie handelt, erhellt sich z. B. daraus, daß die Frege—Whitehead—Russelsche Definition der Kardinalzahlen (ebenso, wie die zweite Cantorsche Definition, *Math. Annalen* 46, S. 481) als nicht befriedigend verworfen werden, während die erste Cantorsche Definition in der Form: Kardinalzahl einer Menge M ist M und jede zu M äquivalente Menge, übernommen wird. An solchen Stellen ist es dem Referenten nicht ganz klar, was vom eingenommenen Standpunkt aus erlaubt ist und was nicht.

Bei der Besprechung des Kontinuumproblems wird die Gödelsche Entdeckung der Unwiderlegbarkeit der Cantorschen Vermutung (in den üblichen, als widerspruchsfrei vorausgesetzten, Axiomensystemen) gar nicht erwähnt; offenbar darum, weil zur Zeit der Abschließung nur eine vorläufige Mitteilung darüber erschienen ist.

L. Kalmár.

Heft 4, Teil 1, ausgegeben im Dezember 1939.

9. W. MAGNUS, *Allgemeine Gruppentheorie*, 51 S.

Es wird über den folgenden Problemkreis berichtet: A. Allgemeine Begriffsbildungen der Gruppentheorie. B. Struktur der Gruppen mit endlichen Untergruppenketten. C. Endliche Gruppen. D. Konstruktion von Gruppen, unendliche Gruppen.

Die Theorie der topologischen und Lieschen Gruppen wird nur berührt, da mit der Einführung des Stetigskeitsbegriffes in Gruppen schon der Bereich der allgemeinen Gruppentheorie verlassen wird.

Es ist zu bedauern, daß für diese wichtige und ausgedehnte Theorie nur 51 Seiten zur Verfügung gestellt werden konnten. Man erhält so zwar eine gute Übersicht über die verschiedensten Begriffsbildungen, Probleme und Ergebnisse der Gruppentheorie, sowie einen ausgezeichneten Wegweiser in der ausgedehnten Literatur, es bleibt aber dem Verfasser kein Raum übrig, die historische Entwicklung oder die Beweismethoden darzustellen, ja auch nur zu berühren.

B. v. Sz. Nagy.

Heft 5, ausgegeben im Dezember 1939.

11. W. KRULL, *Allgemeine Modul-, Ring- und Idealtheorie,* 54 S.

Gemäß der geschichtlichen Entwicklung wird die moderne Theorie der Ringe und Ideale nach zwei Richtungen geordnet, das sind die „additive" (KRONECKER—KÖNIG—LASKER—MACAULAY) und die „multiplikative" (DEDEKIND). Die additive Theorie legt den Untersuchungen einen Ring zugrunde, der auch Nullteiler haben darf. Eine führende Rolle spielen die Prim- und Primärideale. Ausführlich ist die Theorie für Ringe mit Maximalbedingung entwickelt, wobei es meistens auf Durchschnitts-, Summen- und Kettenbildungen ankommt. Hierher gehören die Krullsche Dimensionstheorie der Ringe, die Übertragung der Dedekind-Hilbertschen Primidealtheorie und der Diskriminantansätze auf den Fall der Ringerweiterung. Die multiplikative Idealtheorie legt einen Integritätsbereich mit seinem Quotienkörper zugrunde. Dedekinds Theorie folgend spielt die Produktzerlegung der Ideale in Primideale die Hauptrolle. Hierbei kommt die Bewertungsarithmetik zur Geltung. Das Schlußkapitel ist eine Verschmelzung beider Theorien, angewendet auf Integritätsbereiche, die Ringe mit Minimalbedingung sind. Wieder werden Ringerweiterungen betrachtet, und hier entstehen die Führer-, Grad-, Normen- und Differentensätze.

12. W. KRULL, *Theorie der Polynomideale und Eliminationstheorie,* 53 S.

Als Polynome werden solche von endlich vielen Variabeln über einem Koeffizientenkörper zugelassen. In die Theorie der Polynomideale spielt die allgemeine Idealtheorie nur in beschränkten Maße hinein. Im ersten Kapitel wird die Nullstellen- und Dimensionstheorie betrachtet. Die Hauptpunkte sind hier die Basis- und Nullstellensätze von HILBERT und von LASKER, dann die Nullstellentheorie van der Waerdens und die Dimensionstheorie von KRULL angewendet auf Polynomideale. Das zweite Kapitel bringt die Eliminationstheorie nach den Methoden von KRONECKER und von HENTZELT. Es werden der homogene und nichthomogene Fall einzeln ausgeführt, darunter als Beispiel der lineare Fall, mit Einschluß der Sätze über Trägheitsformen, Resultanten und Diskriminanten. Auch wird hier van der Waerdens Vielfachheitstheorie wiedergegeben. Das dritte Kapitel (Weiterer Ausbau der Polynomidealtheorie) ist verschiedenen Untersuchungen gewidmet, die teils aus der neueren Zeit stammen. Hierüber mögen die Überschriften der einzelnen Paragraphen unterrichten: Lineare Gleichungen im Polynomring. H-Ideale. Die Syzigienkette. Die Hilbertsche Funktion. Das inverse System. Perfekte Ideale. (Die letzen zwei Begriffe rühren von MACAULAY her.) Funktionaldeterminanten und Polynomideale. Polynomideale und Singularitäten algebraischer Gebilde.

13. H. HERMES und G. KÖTHE, *Theorie der Verbände*, 28 S.

Die Verbände sind mit zwei „dualen" Operationen versehene Mengen. Ihre von DEDEKIND („Dualgruppen") geschaffene Theorie hat in den letzten Jahren einen neuen Aufschwung erhalten mit Ausbreitung des Anwendungsgebiets. In den modularen (d. h. Dedekindschen) Verbänden gelten der Satz von JORDAN-HÖLDER und verschiedene Zerlegungssätze. Eine weitere wichtige Klasse bilden die distributiven, insbesondere die Booleschen Verbände. Diese lassen sich als Mengenverbände darstellen und stehen in Beziehung zur Topologie und zur Logik. Die komplementären modularen Verbände von MENGER stehen mit dem projektiven Verknüpfungsaxiomensystem in Verbindung. Es wird noch über die kontinuierlichdimensionale Geometrie von J. NEUMANN und die Darstellung genannter Verbände durch Hauptidealverbände berichtet, auch auf Anwendungen in der Gleichungstheorie und Zahlentheorie hingewiesen.

L. Rédei.

Lothar Heffter, Grundlagen und analytischer Aufbau der Projektiven, Euklidischen, Nichteuklidischen Geometrie, VIII + 197 S. + 2 S. Anhang + 2 S. Berichtigungen und Bemerkungen zum „Lehrbuch der Analytischen Geometrie" des Verfassers, Leipzig und Berlin, B. G. Teubner, 1940.

Das vorliegende Werk zerfällt in 5 Abschnitte. Der erste Abschnitt (*A*) bringt die axiomatische Fundierung der projektiven, affinen und äquiformen Geometrie. Neben der axiomatischen Charakterisierung wird auch F. Kleins gruppentheoretischer Aufbau auseinandergesetzt. Sodann wird schon hier auf die Möglichkeit der Charakterisierung der beiden nichteuklidischen Geometrien hingewiesen. Abschnitte (*B*), (*C*) und (*D*) behandeln die projektive, parallele und äquiforme Geometrie. In diesen drei Abschnitten werden die entsprechenen Geometrien in Punktreihen, in der Ebene und im Raum entwickelt. Die der Betrachtung zugrunde liegenden Gebilde sind die linearen bzw. diejenigen zweiter Ordnung. Gearbeitet wird mit Koordinaten. Nach den Betrachtungen in (*A*) werden gleich zu Beginn von (*B*) projektive Koordinaten eingeführt. In (*C*) und (*D*) werden entsprechend spezielle affine (Hessesche) bzw. äquiforme Koordinaten eingeführt. In (*D*) wird das allgemeine Prinzip der projektiven Maßbestimmung auseinandergesetzt, das sich bei Zugrundelegung eines absoluten Gebildes ergibt. Abschnitt (*E*) behandelt die hyperbolische bzw. elliptische Geometrie. Verfasser charakterisiert dieselbe in wohlbekannter Weise dadurch, daß er im projektiven Raum eine Fläche zweiten Grades als absolutes Gebilde auszeichnet. Dargestellt werden: Bewegungslehre, Grundlagen der metrischen Geometrie (Cayley—Kleinsche Größenlehre), Cliffordsche

Fläche, verschiedene Modelle der nichteuklidischen Geometrien. In einem Anhang wird eine geänderte Einführung der Kleinschen Modelle behandelt.

Des Buch stellt einen recht systematischen Aufbau der euklidischen und nichteuklidischen Geometrien dar, ausgehend von der projektiven Geometrie.

O. Varga.

Später zu besprechen. — A être analysés plus tard. — To be reviewed later on.

Marston Morse, The Calculus of Variations in the Large (American Math. Society Colloquium Publications, Vol. XVIII), IX + 368 pages, New York, American Mathematical Society, 1934.

J. L. Walsh, Interpolation and Approximation by Rational Functions in the Complex Domain (American Math. Society Colloquium Publications, Vol. XX), IX + 382 pages, New York, American Mathematical Society, 1935.

C. Carathéodory, Geometrische Optik (Ergebnisse der Math. und ihrer Grenzgebiete, 4. Bd., H. 5), IV + 104 S., Berlin, J. Springer, 1937.

R. Courant und D. Hilbert, Methoden der mathematischen Physik, II. Band (Grundlehren der math. Wissenschaften, Bd. XLVIII), XVI + 549 S., Berlin, J. Springer, 1937.

Maurice Fréchet, Méthode des fonctions arbitraires, théorie des événements en chaine dans le cas d'un nombre fini d'états possibles (Traité du Calcul des Probabilités par EMILE BOREL, t. I, fasc. III, 2. livre), XIII + 315 pages, Paris, Gauthier-Villars, 1938.

Charles N. Moore, Summable Series and Convergence Factors (American Math. Society Colloquium Publications, Vol. XXII), VI + 105 pages, New York, American Mathematical Society, 1938.

Louis de Broglie, La Mécanique Ondulatoire des systèmes de corpuscules (Collection de Physique Math., fasc. V), VI + 223 pages, Paris, Gauthier-Villars, 1939.

Jacques Solomon, Protons, neutrons, neutrinos (Collection de Physique Math., fasc. VI), XIII + 228 pages, Paris, Gauthier-Villars, 1939.

J. Tinbergen, Une méthode et son application au mouvement des investissements (Verification statistique des théories des cycles économiques, I), 178 pages, Genève, Société des Nations, Services d'études économiques, 1939.

Francis Perrin, Mécanique Statistique Quantique (Traité du Calcul des Probabilités par EMILE BOREL, t. II, fasc. V), III + 224 pages, Paris, Gauthier-Villars, 1939.

Louis Bachelier, Les nouvelles méthodes du Calcul des Probabilités, VIII + 71 pages, Paris, Gauthier-Villars, 1939.

Jean Ville, Etude critique de la notion de collectif (Monographies des Probabilités, publiées sous la direction de M. EMILE BOREL, fasc. III), I + 144 pages, Paris, Gauthier-Villars, 1939.

A. Adrian Albert, Structure of Algebras (American Math. Society Colloquium Publications, Vol. XXIV), XI + 210 pages, New York, American Mathematical Society, 1939.

Emile Sevin, Physique stellaire, essai de synthèse (Extrait de Bulletin Astronomique), 80 pages, Paris, Gauthier-Villars, 1939.

Tsurusaburo Takasu, Differentialgeometrien in den Kugelräumen, Bd. II: Laguerresche Differentialkugelgeometrie, XX + 444 S., Tokyo, Maruzen Co. Ltd., 1939.

L. Cagniard, Réflection et réfraction des ondes séismiques progressives, XI + 255 pages, Paris, Gauthier-Villars, 1939.

Kurt Reidemeister, Die Arithmetik der Griechen (Hamburger math. Einzelschriften, 26. H.), 31 S., Leipzig und Berlin, B. G. Teubner, 1940.

C. Carathéodory, Elementare Theorie des Spiegelteleskops von B. Schmidt (Hamburger math. Einzelschriften, 28. H.), IV + 36 S., Leipzig und Berlin, B. G. Teubner, 1940.

W. Blaschke, Mathematik und Leben (Hamburger math. Einzelschriften, 27. H.), 13 S., Leipzig und Berlin, B. G. Teubner, 1940.

Fritz Emde, Tafeln elementarer Funktionen, XII + 181 S., Leipzig und Berlin, B. G. Teubner, 1940.

Wolfgang Gröbner, Idealtheoretischer Aufbau der algebraischen Geometrie, Teil I (Hamburger math. Einzelschriften, 30. H.), V + 56 S., Leipzig und Berlin, B. G. Teubner, 1941.

Wilhelm Blaschke, Nicht-Euklidische Geometrie und Mechanik, I, II, III (Hamburger math. Einzelschriften, 34. H.), III + 82 S., Leipzig und Berlin, B. G. Teubner, 1942.

Kurt Reidemeister, Mathematik und Logik bei Plato (Hamburger math. Einzelschriften, 35. H.), I + 20 S., Leipzig und Berlin, B. G. Teubner, 1942.

Béla v. Sz. Nagy, Spektraldarstellung linearer Transformationen des Hilbertschen Raumes (Ergebnisse der Math. und ihrer Grenzgebiete, 5. Bd., H. 5), IV + 80 S., Berlin, Springer-Verlag, 1942.

H. Tietze, Ein Kapitel Topologie (Hamburger math. Einzelschriften, 36. H.), VII + 47 S., Leipzig und Berlin, B. G. Teubner, 1942.